

Aus dem Fachbereich 06 Psychologie
der Justus Liebig-Universität Giessen
(Dekan: Prof. Dr. R. König)

**EINE EMPIRISCHE UNTERSUCHUNG
ZU URTEILSPROZESSEN
UND EXPLORATIONSVERHALTEN
VON PSYCHIATERN**

Inaugural-Dissertation
zur Erlangung des Doktorgrades
des Fachbereichs 06 Psychologie
der Philosophischen Fakultät
der Justus Liebig Universität Giessen

vorgelegt von
Reinhard Leichner
Bielitz

Giessen 1972

Aus dem Fachbereich 06 Psychologie
der Justus Liebig-Universität Giessen
(Dekan: Prof. Dr. R. König)

**EINE EMPIRISCHE UNTERSUCHUNG
ZU URTEILSPROZESSEN
UND EXPLORATIONSVERHALTEN
VON PSYCHIATERN**

Inaugural-Dissertation
zur Erlangung des Doktorgrades
des Fachbereichs 06 Psychologie
der Philosophischen Fakultät
der Justus Liebig Universität Giessen

vorgelegt von
Reinhard Leichner
Bielitz

Giessen 1972



D 26

Dekan: Prof. Dr. R. König

1. Berichterstatter: Prof. Dr. A. Spitznagel

2. Berichterstatter: Prof. Dr. H. Hartmann

Tag der mündlichen Prüfung:

12. Mai 1972

Vorwort

Die Untersuchung wurde an der neuropsychiatrischen Klinik der Justus Liebig-Universität in Giessen durchgeführt. Dem Direktor dieser Klinik, Herrn Prof. Dr. A. Derwort, danke ich dafür, mir die Möglichkeit für diese Untersuchung gegeben zu haben.

Mein besonderer Dank gilt denjenigen Ärzten der Klinik, die als Versuchspersonen an der Untersuchung teilnahmen und dafür sehr viel Zeit opferten. Ihre Namen bleiben wegen der zugesicherten Anonymität ungenannt.

Herr Prof. Dr. A. Spitznagel hat mich durch kritische Anmerkungen und durch Vorschläge bei der Planung und Darstellung der Untersuchung unterstützt, wofür ich ihm sehr dankbar bin.

I. Allgemeines	1
1.1. Physiologischer Unfall	1
1.2. Differenzierung der Unfälle	2
1.3. Ursachen	3
1.4. Symptome	4
1.5. Diagnostik	5
1.6. Prognose	6
II. Psychologisches Umfeld	7
2.1. Definition	7
2.2. Symptome	8
2.3. Diagnostik	9
2.4. Prognose	10
III. Differenzielle Diagnostik	11
3.1. Psychologischer Unfall	11
3.2. Epilepsie	12
3.3. Parkinson	13
3.4. Demenz	14

Inhaltsverzeichnis

	Seite
A : EINLEITUNG	1
1. Psychiatrisches Urteilen als Diagnostizieren	1
2. Psychologisches Urteilen	5
B : FRAGESTELLUNG	13
C : UNTERSUCHUNG	18
1. Versuchsplan und Versuchsdurchführung	18
2. Beurteilungsinstrument : Polaritäten-Profil	23
3. Versuchspersonen	24
31. Psychiater	24
32. Patienten	24
4. Methode	24
D : ERGEBNISSE	33
I Beurteilung der Patienten	33
1. Allgemeine Aspekte	34
11. Psychologisches Urteil	34
111. Differenziertheit der Urteile	34
112. Skalen	37
113. Patienten	41
114. Beurteiler	49
115. Kennmatrizen	50
12. Psychiatrisches Urteil	53
121. Stabilität	55
122. Validität	60
2. Differentielle Aspekte	67
21. Psychologisches Urteil	67
211. Differenziertheit der Urteile	67
212. Skalen	68
213. Patienten	69
214. Urteile	73

	Seite
22. Psychiatrisches Urteil	77
221. Stabilität	77
2211. Intraindividuelle Unterschiede : Psychiater	77
2212. Interindividuelle Unterschiede : Psychiater	78
2213. Differentielle Aspekte : Patienten	79
222. Validität	80
2221. Intraindividuelle Unterschiede : Psychiater	80
2222. Interindividuelle Unterschiede : Psychiater	82
2223. Differentielle Aspekte : Patienten	85
II Psychologische Beurteilung der Psychiater	88
1. Ärzte als Beurteiler	90
11. Bestimmung der Urteile durch Diagnosegruppen und Explorationssituation	90
111. Differenziertheit der Urteile	90
112. Skalen	92
113. Beurteilte	94
114. Beurteilungen	95
12. Bestimmung der Urteile durch die individuellen Patienten	99
2. Psychiater als Beurteilte	100
21. Differenziertheit der Urteile	101
22. Skalen	102
23. Beurteilungen	104
24. Beurteiler	105
E : ZUSAMMENFASSUNG, DISKUSSION, KRITIK	106
F : SCHLUSSBEMERKUNGEN	118
Literaturverzeichnis	122
Anhang	127

A. EINLEITUNG

1) Psychiatrisches Urteilen als Diagnostizieren

Psychiatrisches Diagnostizieren impliziert einen psychischen Prozeß, der grob in Informationserhebung über Exploration der Patienten, Informationsaufnahme, Informationsverarbeitung und Abgabe eines psychiatrischen Urteils gegliedert werden kann. Dieser Prozeß läuft bei der diagnostischen Beurteilung eines Patienten nicht nur einmal ab, vielmehr handelt es sich - etwa im Sinne des kybernetischen Modells - um ein mehrmaliges Durchkreisen der vier genannten Prozeßstadien, wobei jeweils in Abhängigkeit vom erreichten Urteil spezifische neue Information erhoben wird. Hypothesen über den Patienten werden formuliert, zur Überprüfung der Hypothesen spezifische Information erhoben, die Information wird aufgenommen, verarbeitet und die Hypothese angenommen oder verworfen. Die Hypothesen, die der Psychiater an die Patienten heranträgt, sind Diagnosen als Klassen mit bestimmten sie definierenden Symptomen. Das Vorhandensein dieser Symptome wird am Patienten im schleifenförmigen diagnostischen Prozeß überprüft. Nach mehrmaligen Überprüfungen, die zu Zwischen-Diagnosen führen, erfolgt eine End-Diagnose, d.h. der Patient wird endgültig in eine der Krankheitsgruppen kategorisiert.

Die Exploration unterliegt der Steuerung durch die Hypothesen, d.h. der Psychiater setzt in Abhängigkeit vom Ausmaß des Wissens über die Patienten mit seiner Exploration mehr oder minder gezielt an bestimmten noch unklaren psychischen Bereichen an. Diese diagnostische Situation wiederholt sich fortlaufend während einer einzigen Exploration. Sie ist aber auch dann gegeben, wenn die Exploration abgebrochen und später wieder aufgenommen wird. Dieser Sachverhalt trifft gewöhnlich bei der stationären Aufnahme eines Patienten zu; der Patient wird meist schon mit einer vorläufigen Diagnose und Beschreibung in die Station eingeliefert. In der Poliklinik befindet sich der Psychiater dagegen häufig in einer anderen Situation. Er wird von Personen aufgesucht, die ihm unbekannt sind und nicht von einem anderen Arzt mit bestimmten diagnostischen Vermutungen überwiesen wurden. Hier

werden zunächst allgemeine Bemerkungen und Schilderungen des Patienten über seine Schwierigkeiten sowie sein mimisches und motorisches Verhalten zur Entstehung der ersten diagnostischen Vermutungen des Psychiaters beitragen. Auch wird der Psychiater durch Kenntnis der Auftretenshäufigkeit verschiedener Krankheitsgruppen nicht völlig unvoreingenommen eine Exploration beginnen. Dennoch ist er meist gezwungen, sämtliche zur Diagnose relevanten Bereiche abzufragen. Die Gesprächsabfolge orientiert sich gewöhnlich nach der ersten Kontaktaufnahme an folgendem Schema: Zunächst wird nach organischen Erkrankungen gefragt, dann eine Schilderung des sozialen und beruflichen Werdegangs vom Patienten verlangt. Nach dieser eher allgemeinen Information beginnt dann schließlich die spezifische, auf Krankheitssymptome abzielende Exploration.

Die drei Bereiche Kontaktaufnahme, biographische und soziale Anamnese dienen gewissermaßen als vorläufige diagnostische Zwischenstation, die die Grenzen des Problembereichs, d.h. die mögliche Erkrankung, in etwa abstecken helfen. Für die psychiatrische Diagnose gewöhnlich am wichtigsten ist die spezifische Exploration zur Erkrankung. Möglich ist eine exakte psychiatrische Diagnose auch allein aufgrund der Information aus den drei genannten Bereichen, wobei jedoch organische Erkrankungen hier unberücksichtigt bleiben sollen, da zu deren Diagnose die biographische Anamnese meist ausreicht. Prinzipiell dürfte sich jedoch der psychiatrisch-diagnostische Prozeß aufgrund der Information aus den drei Bereichen Kontaktaufnahme, biographische und soziale Anamnese und der Information aus der spezifischen Exploration zur Erkrankung unterscheiden. Motorische und mimische Verhaltensweisen mit Symptombedeutung finden bei diesen Überlegungen jedoch keine Berücksichtigung:

In der spezifischen Exploration wird das Vorhandensein von Symptomen direkt überprüft. Symptome werden dabei jedoch nicht unmittelbar erfragt; vielmehr zielen die Fragen geringfügig an den Symptomen vorbei, um die verbalen Reaktionen der Patienten anzuregen. Die vom Patienten gelieferte Information bedarf im Idealfall keiner interpretativen Weiterverarbeitung; der Information ist direkt zu entnehmen, ob bestimmte Symptome vorliegen oder nicht. Manchmal, vor allem bei mißtrauischen, wenig auskunftsbereiten Patienten müssen die Aussagen interpretativ aufbereitet werden. Im Gegensatz zu diesem diagnostischen Vorgehen muß die Information aus den

Bereichen Kontaktaufnahme, biographische und soziale Anamnese in weitaus stärkerem Maße einer Interpretation unterzogen werden, d.h. aufgrund der Angaben des Patienten werden mögliche Symptome erschlossen, die Symptome werden nicht direkt überprüft. Ausnahme davon sind lediglich diejenigen Angaben von Patienten, die unmittelbar Symptombedeutung besitzen. Z.B. gilt ein in der sozialen Anamnese berichteter plötzlicher beruflicher Leistungsabfall als Symptom für eine Hebephrenie. Im allgemeinen müssen jedoch die Angaben auf ihre Symptomrelevanz hin interpretiert werden.

Der dargestellte Unterschied zwischen den beiden diagnostischen Prozeßformen ist - wie schon angedeutet - kein qualitativer. In beiden Fällen kann interpretiert bzw. die erhaltene Information direkt zur Überprüfung der Hypothesen verwendet werden. Quantitativ liegen jedoch sicherlich insofern Unterschiede vor, als das Ausmaß der interpretativen Verarbeitung der Information aus den Bereichen Kontaktaufnahme, biographische und soziale Anamnese größer ist als das der Information aus der spezifischen Exploration zur Erkrankung.

Beide der hier beschriebenen psychiatrisch-diagnostischen Prozeßformen führen zu einer psychiatrischen Diagnose, d.h. der Patient wird in eine der psychiatrischen Krankheitsgruppen kategorisiert und damit auch psychiatrisch näher beschrieben.

Psychiatrische Krankheitsgruppen sind - zumindest in der klassischen Psychiatrie - logische, d.h. konjunktive und disjunktive Klassen. Beim psychiatrischen Entscheidungsvorgang, der zur Kategorisierung führt, wird also - wie schon beschrieben - ermittelt, ob Merkmale am zu beurteilenden Patienten vorhanden bzw. nicht vorhanden sind und entsprechend wird eine Kategorisierung vorgenommen. Beim Ermittlungsvorgang spielt die Interpretation der Information eine mehr oder minder große Rolle. Die Merkmale als Symptome können Charakteristika größerer oder geringerer zeitlicher Erstreckung sein und direkt beobachtbar, d.h. durch Exploration direkt erfahrbare oder verdeckte sein. Verdeckte Charakteristika müssen aufgrund anderer, beobachtbarer, erschlossen, d.h. interpretativ erfaßt werden.

Untersuchungen, die sich mit dem psychiatrischen Diagnostizieren beschäftigen, überprüfen gewöhnlich Reliabilität und Validität des Diagnostizierens (vgl. RANABAUER, 1968; ZUBIN, 1967). In der letzten Zeit werden insbesondere Untersuchungen

durchgeführt, in denen der Einfluß von verschiedenen Faktoren wie Erfahrung, soziale Haltung des Psychiaters usw. auf Reliabilität, Validität und Auftretenshäufigkeit von Diagnosen ermittelt wird. Nur kurz sei erwähnt, daß die Reliabilität und Validität des Diagnostizierens im allgemeinen sehr gering ist, wobei sich jedoch Unterschiede insbesondere in Abhängigkeit von Krankheitsgruppen ergeben. Schizophrene werden im allgemeinen reliabler und valider beurteilt als andere Patientengruppen. Weiterhin sei darauf hingewiesen, daß bezüglich Reliabilität und Validität des psychiatrischen Diagnostizierens erhebliche interindividuelle Unterschiede zwischen den beurteilenden Psychiatern auftreten.

Aufgrund der verschiedenen Untersuchungen wird die Kritik am psychiatrischen nosologischen System sowie an der psychiatrischen diagnostischen Situation immer größer. So konnten WARD und Mitarbeiter (1962) nachweisen, daß 37% des Mangels an Übereinstimmung zweier Diagnostiker Unterschiede und Irrtümer bei der Explorationstechnik zugeschrieben werden kann, während 58% der Unterschiede auf die Inadäquatheit der Nomenklatur zurückzuführen sind.

Die Kritik am nosologischen System betrifft vorrangig zweierlei Aspekte:

- 1) Die Diagnosegruppen sind durch Symptome definiert, die es in der Realität gar nicht gibt. Der Psychiater hat somit die Patienten in inhaltlich nicht angemessene Diagnosegruppen zu kategorisieren.
- 2) Die Form der Diagnosegruppen, nämlich logische Klassen, sind wegen der Dichotomie der Merkmale nicht realitätsangemessen. Darüber hinaus implizieren disjunktive Klassen - und die größte Anzahl psychiatrischer Diagnosegruppen sind disjunktive Klassen - selbst eine der größten Fehlerquellen. Zur Kategorisierung in disjunktive Klassen müssen nämlich nicht alle die Klasse definierenden Merkmale am Patienten feststellbar sein. Da Merkmale verschiedener Klassen z.T. identisch sind, ist die Kategorisierung in eine der Klassen außerordentlich schwierig. Es bedarf daher häufig sehr spezifischer Information, um differentialdiagnostisch Patienten in eine von zwei sich z.T. überschneidenden disjunktiven Klassen einzuordnen.

Kritiker wie REDLICH und FREEDMAN (1966) fordern daher die Abschaffung der Diagnosekategorien und verlangen stattdessen die Beschreibung von Verhaltensstörungen mit ihren motivationalen Hintergründen. Es handelt sich dabei um eine Aufgabe, der gewöhnlich der Klinische Psychologe nachkommt.

Andere Autoren versuchen dagegen, neue Diagnosesysteme zu erstellen. Z.B. bemüht sich LORR (1963) darum, für Psychosen ein neues Diagnosesystem über Beurteilung anhand von Skalen zu schaffen. Diese Diagnosesysteme sollen realitätsangemessen sein, d.h. den tatsächlichen Typen von Verhaltensstörungen entsprechen. Außerdem bestehen solche Diagnosesysteme gewöhnlich aus relationalen Klassen; relationalen Klassen liegen dieselben Merkmale oder Dimensionen zugrunde, die Klassen sind durch unterschiedliche Ausprägung auf den Dimensionen definiert. Solche relationalen Klassen werden gewöhnlich innerhalb psychologischer Diagnosesysteme verwendet. Die entsprechenden Ansätze innerhalb der Psychiatrie weisen auf, daß die so ermittelten Patiententypen häufig nur wenig mit Diagnosegruppen aus dem Psychiatrischen Diagnosesystem übereinstimmen.

2) Psychologisches Urteilen

Praktisches psychologisches Diagnostizieren bedeutet nicht nur - wie psychiatrisches Diagnostizieren - Kategorisieren von Patienten. Vielmehr verlangt es auch Deskription von Verhaltensstörungen und Versuche, die Störungen psychologisch zu erklären. Die Deskription kommt einer Kategorisierung jedoch dann nahe, wenn zur Beschreibung verschiedener Patiententypen jeweils dieselben Dimensionen benutzt werden, was häufig sicherlich zutrifft (vgl. KELLY, 1955). Eindeutig ist diese Situation gegeben, wenn man - wie das in Untersuchungen zur klinischen Urteilsbildung üblich ist - bei der Beurteilung der Patienten jeweils dieselben Skalen benutzt. Der Urteiler hat gebotene Information aufzunehmen, zu verarbeiten und sein Urteil anhand der Personen charakterisierenden Skalen abzugeben. Die Interpretation der Information impliziert diesen Urteilsprozeß, wenn die Beurteilungsskalen Eigenschaften beinhalten, die nicht direkt beobachtbar sind. Eine statistisch vorgenommene Typisierung der anhand der Skalen beurteilten Patienten führt gewöhnlich zu relationalen Klassen, die Dimensionierung der Beurteilungsskalen zu Beurteilungsdimensionen, mit denen die Patienten auf Dimensionsebene näher charakterisiert werden.

Dieser Beurteilungsvorgang anhand von Skalen unterscheidet sich vom psychiatrischen Diagnostizieren insofern, als dem Urteiler Patiententypen sowie Skalendimensionen nicht klar bewußt sein müssen. Patiententypen und Skalendimensionen liefert erst die statistische Bearbeitung. Im Gegensatz dazu muß der Psychiater von den psychiatrischen Klas-

sen eine weitaus klarere Vorstellung besitzen, da er eine direkte Kategorisierung in die Klassen vornimmt. Die die Klassen definierenden Symptome müssen möglichst eindeutig bekannt sein.

Ein den psychologisch-diagnostischen Prozeß beschreibendes Modell haben SARBIN et al. (1960) entwickelt. An diesem Modell haben sich die oben genannten Überlegungen in Bezug auf Patiententypen und Beurteilungsdimensionen orientiert. Dieses Modell basiert vorrangig auf dem Ansatz von BRUNSWIK (1956, 1955) und KELLY (1955).

Die Autoren stellen - im Sinne Brunswiks - die Umwelt, z.B. Personen, und das Ereignisse in dieser Umwelt beurteilende Individuum gegenüber. Der Umwelt liegen Dimensionen zugrunde, d.h. die zu beurteilenden Personen weisen Eigenschaften als Dimensionen auf. Diesen Dimensionen mehr oder minder entsprechende Dimensionen hat der Urteiler durch seine Erfahrung mit der Umwelt - im Sinne eines sozialen Lernprozesses - aufgebaut. Anhand dieser Dimensionen beschreibt der Urteiler Personen, indem er sie innerhalb dieses kognitiven Raums lokalisiert, d.h. in relationale Klassen einordnet. Das Modell geht also davon aus, daß die Beurteilung nicht nur anhand einzelner Skalen, sondern anhand globaler Urteilsdimensionen erfolgt, und entsprechend impliziert dieses Modell Patiententypen.

Die Beziehung der zur Beurteilung benutzten Skalen innerhalb einer Dimension sowie das Gefüge der Dimensionen zueinander, der kognitive Raum, wird seit ASCH (1946) und insbesondere seit CRONBACH (1955, 1958) als ein Aspekt impliziter Persönlichkeitstheorien bezeichnet. Der Begriff implizite Persönlichkeitstheorie weist schon darauf hin, daß die exakte Dimensionierung nicht unmittelbar gegeben, sondern erst nach Beurteilung anhand der Skalen ermittelbar ist. Derselbe Sachverhalt trifft für die Patiententypen zu. Es ist natürlich möglich, vor allem einen Experten nach den von ihm benutzten Urteilsdimensionen und Patiententypen direkt zu befragen. Eine entsprechende Auskunft unterliegt aber sicherlich starken Verzerrungen, da der psychologisch-diagnostische Prozeß auch beim Experten auf einer nicht sehr hohen Bewußtseinssebene abläuft. Daher ist die Ermittlung von Urteilsdimensionen und Patiententypen über den realen diagnostischen Vorgang, nämlich Beurteilung von Patienten anhand von Skalen, sinnvoller.

Psychologisch besonders interessant ist der Kategorisierungsvorgang als Entscheidungs-

prozeß, für den die Autoren verschiedene modellhafte Möglichkeiten anbieten. Eine exakte Analyse des Kategorisierungsvorgangs ist bisher weder bezüglich psychologischen noch psychiatrischen Urteilens gelungen. Seit Ende der fünfziger und insbesondere Anfang der sechziger Jahre ist jedoch bezüglich psychologischen Urteilens eine Flut von Untersuchungen erschienen, in denen der Einfluß von verschiedenen Faktoren auf den Urteilsprozeß und damit auf das Urteil überprüft wird. Dabei ist interessant festzustellen, daß meist erst neuere Untersuchungen der Kompliziertheit des Prozesses insofern gerecht werden, als sie mehrere Faktoren in ihre Untersuchungspläne aufnehmen und somit auch die Interaktion der Faktoren überprüfen. Aufgrund der Kenntnis des Einflusses der Faktoren wird häufig auf Aspekte des Prozesses selbst zurückgeschlossen. Weiterhin wird der Urteilsprozeß mit logischen und statistischen Modellen simuliert, um so zu einem näheren Verständnis des Prozesses zu gelangen.

Als Faktoren, die Einfluß auf den Prozeß und damit auf das Urteil nehmen können, sind bisher - grob gegliedert - bekannt:

- 1a) Art der beobachteten Stimuli als Information
- 1b) Umfang der dargebotenen Information
- 1c) Reihenfolge der Informationsdarbietung
- 2) Kognitive und affektive Komponenten der Person, an der die Stimuli beobachtet wurden
- 3) Kognitive, affektive und motivationale Komponenten des Urteilers
- 4) Interaktion zwischen Urteiler und Beurteiltem
- 5) Insbesondere im klinischen Bereich das Wissen des Urteilers von der Auftretenshäufigkeit bestimmter Syndrome oder auch Patientengruppen
- 6) Die Kenntnis sozialer Folgen von diagnostischen Entscheidungen auf diese Entscheidungen.

Eine Übersicht über die verschiedenen Untersuchungen geben u.a. TAGIURI (1969) und FRANK (1970).

Im folgenden sollen einige Untersuchungen referiert werden, die den Urteilsprozeß mehr oder minder direkt berühren. Es handelt sich hauptsächlich um Untersuchungen zur impliziten Persönlichkeitstheorie. Seit BRUNER und TAGIURI (1954) und CRONBACH (1955) werden drei Komponenten der impliziten Persönlichkeitstheorie ausgegliedert:

- 1) die Tendenz, bestimmte Eigenschaften konsistent höher oder niedriger einzustufen,
- 2) die Tendenz, bestimmte Eigenschaften extrem oder neutral einzuschätzen und

3) die Tendenz, bestimmte Eigenschaften als miteinander verbunden oder unabhängig zu sehen.

Im Mittelpunkt der bisherigen sowie der folgenden Überlegungen steht die dritte Komponente, die Beziehung zwischen Eigenschaften.

Zur Ermittlung des Zusammenhangs zwischen Eigenschaften finden hauptsächlich zwei Methoden Anwendung:

- 1) Interkorrelation der zur Beurteilung verwendeten Skalen bzw. Eigenschaften (CRONBACH)
- 2) Angabe der Wahrscheinlichkeit des Auftretens verschiedener Eigenschaften bei Vorliegen einer oder mehrerer anderer Eigenschaften (unabhängig von real Beurteilten). Dieser Ansatz geht auf BRUNER, SHAPIRO und TAGIURI (1958) zurück. Er wurde von HAYS (1958) methodisch ausgearbeitet.

Methode 2) kommt dem Prozeßcharakter des Urteilens insofern näher, als beim psychologischen Urteilen tatsächlich im Rahmen der impliziten Theorie Eigenschaften aufgrund anderer - unabhängig von real Beurteilten - erschlossen werden. Dieser Aspekt wird mit Methode 1) weniger eindeutig erfaßt, obgleich sich der Urteiler beim Einschätzen einer Person anhand von Skalen bzw. Eigenschaften über die Abhängigkeit einiger Skalen bzw. Eigenschaften voneinander z.T. klar sein dürfte. TODD und RAPPAPORT (1964) verglichen die beiden Beziehungsmethoden und stellten eine recht gute Übereinstimmung in der relativen Größe der Beziehungskoeffizienten fest. Dagegen war die Kongruenz nach Dimensionierung weniger hoch. Zur Dimensionierung der nach der Hays'schen Methode gewonnenen Koeffizientenmatrix wurde eine spezielle Art der multidimensionalen Skalierung verwendet, im anderen Fall wurde eine Faktorenanalyse durchgeführt. Wie weit die Unterschiede in der Struktur auf die Dimensionierungsmethoden zurückzuführen sind, ist dabei nicht klar.

Was die mit den beiden genannten multidimensionalen statistischen Verfahren oder anderen ähnlichen statistischen Verfahren wie z.B. der Vektormethode von TUCKER und MESSICK (1963) gewonnenen Faktoren tatsächlich repräsentieren, ist immer noch eine strittige Frage, d.h. die Bedeutung der Dimensionen und die des Gefüges der Dimensionen als Aspekte der impliziten Persönlichkeitstheorien ist nicht geklärt.

CATTELL (1957) und CATTELL und SCHEIER (1961) meinen, damit seien die - im Sinne Brunswiks - in der Umwelt, d.h. in den beurteilten Personen vorliegenden Eigenschaften erfaßt, womit natürlich schon eine spezifische Bedeutung des Eigenschaftsbegriffes impliziert ist.

HATFIELD und SHAPIRO (1965) stützen diese Ansicht durch eine Untersuchung, in der zumindest bezüglich der ersten varianzstarken Faktoren eine gute Übereinstimmung zwischen Beurteilungsfaktoren und Fragebogenfaktoren errechnet wurde (vgl. auch PETERSON, 1965). Wie weit jedoch die methodische Gemeinsamkeit von Fragebogen und Beurteilung, nämlich in beiden Fällen eine sprachliche Erfassung von Eigenschaften, die Übereinstimmung bedingt, ist dabei nicht geklärt. Entsprechend sind Autoren wie GRAUMANN (1960) und KAMINSKI (1959) der Ansicht, die Dimensionen reflektierten semantische Zusammenhänge, d.h. kognizierte, aufgrund semantischer Übereinstimmung entstandene Dimensionen bestimmten die perzipierten, die bei der Beurteilung realer Personen feststellbaren. Diese Meinung wird u.a. gestützt durch eine Untersuchung von DEROGATIS et al. (1970), in der nachgewiesen wurde, daß die Verknüpfung von Symptombeschreibungen zu 4 Clustern übereinstimmt mit 4 Faktoren, die bei der Beurteilung von Patienten anhand derselben Symptombeschreibungen zustandekommen. Die Vermutung widerlegt KEIL (1970), der einen negativen Zusammenhang zwischen den kognizierten und perzipierten Eigenschaftsbeziehungen fand.

Sinnvoller erscheint eine Mittelstellung bezüglich dieser Frage. Da der Urteiler Eigenschaften nur sprachlich zu differenzieren lernt, und er anhand von Skalen auch nur sprachliche Beurteilungen treffen kann, hat die Sprache, hier die semantische Bedeutung der Eigenschaften, als Erfassungsmodus sicherlich starken Einfluß auf die Urteilsstruktur. Entsprechend kann mit Hilfe der Multitrait-multimethod-Matrix (CAMPBELL und FISKE, 1959) immer wieder die starke Abhängigkeit des Inhalts einer Eigenschaft von ihrer Erfassungsmethode nachgewiesen werden. Da der Einfluß der Sprache auf die Erfassung von Eigenschaften ausschlaggebend Einfluß nimmt, ist auch ein Anteil rein sprachlicher (denotativ, konnotativ, pragmatisch) Beziehung zwischen den beurteilten bzw. zur Beurteilung herangezogenen Eigenschaften zu erwarten. Einen ähnlichen Standpunkt nehmen BRUNER und TAGIURI (1954), HAYS (1958), CRONBACH (1958) und COHEN (1969) ein, die diese Beziehungen als einen Aspekt impliziter Persönlichkeitstheorien betrachten und - unabhängig von der genannten Auseinandersetzung - implizite Persönlichkeitstheorien als eigenständigen Untersuchungsgegenstand für würdig befinden.

Aufgrund des Einflusses der Sprache auf die Beziehungen zwischen den Skalen ist zu erwarten, daß die über vorgelegte Eigenschaftsskalen gewonnenen Urteilsstrukturen relativ stabile Gebilde sind, die sich nur wenig in Abhängigkeit vom Stimulusmaterial

also den Beurteilten, ändern. Größere Unterschiede könnten nur dann erwartet werden, wenn der Urteiler die Eigenschaften zur Beurteilung selbst wählen darf. Eine Untersuchung von DORNBUSCH et al. (1965) legt die Vermutung nahe, daß selbst unter diesen Bedingungen der Einfluß der Beurteilten auf die Urteile relativ gering ist. Gegenteiliger Ansicht sind jedoch BEACH und WERTHEIMER (1961).

NORMAN und GOLDBERG (1966) wiesen durch Simulation mit Hilfe des Monte-Carlo-Verfahrens nach, daß die Urteilsstruktur bei Beurteilung anhand von vorgelegten Eigenschaften fast völlig unabhängig von den Beurteilten ist. In die gleiche Richtung weist auch eine Untersuchung von PASSINI und NORMAN (1966; vgl. auch COHEN, 1969), während WARR und KNAPPER (1968) bei sehr heterogenen Beurteilten gegensätzliche Ergebnisse fanden.

Eindeutigen Einfluß scheinen dagegen unterschiedliche Informationsarten über dieselben Personen auf die Urteilsstrukturen zu nehmen (COHEN, 1969). Entsprechend weist auch GOLDBERG (1966) mit Hilfe der Multitrait-multimethod-Matrix nach, daß das Urteil über Eigenschaften in starkem Maße von Daten verschiedener Testverfahren abhängig ist, womit eines der Kernprobleme der Diagnostik angeschnitten wird.

Aufgrund der Annahmen zur impliziten Persönlichkeitstheorie ist weiterhin zu erwarten, daß die Urteilsstrukturen unterschiedlicher Urteiler beträchtlich differieren, da die Verwendung von Eigenschaftswörtern individuell erlernt wird. Diese Vermutung wird durch eine Reihe von Untersuchungen bestätigt (u.a. CRONBACH, 1958; LOEHLIN, 1967; COHEN, 1967, 1969). Diese interindividuellen Unterschiede können den inhaltlichen Aspekt und den Aspekt der Differenziertheit der Urteilsstrukturen betreffen. Inhaltliche Unterschiede lassen sich z.B. an einer geringen Übereinstimmung der faktoren-analytisch gewonnenen Urteilsstrukturen nachweisen. Unterschiede in der Differenziertheit können wiederum nach faktorenanalytischer Berechnung - anhand unterschiedlicher Varianzstärken des jeweils ersten Faktors, unterschiedlicher Gesamtzahl signifikanter Faktoren sowie unterschiedlicher Größe der extrahierten gemeinsamen Varianz aufgezeigt werden. Inhaltliche Unterschiede der Urteilsstrukturen zeigen an, daß die Urteiler differierende Beurteilungsdimensionen benutzen, was Bedeutungsvarianz der Beurteilungsskalen impliziert. Dieser Sachverhalt stellt ein wichtiges Problem der Persönlichkeitspsychologie und Diagnostik dar.

Interindividuelle Unterschiede im Differenziertheitsgrad kennzeichnen Unterschiede im Ausmaß der kognitiven Komplexität bzw. Einfachheit der Urteiler (CROCKETT, 1965;

BIERI et al., 1966). Der Grad der kognitiven Differenziertheit steht dabei u.a. in positiver Beziehung zum Ausmaß der Informationsaufnahme (SHRAUGER, 1967), der beschreibenden vs. bewertenden Stellungnahme und der differenzierten Verarbeitung widersprüchlicher Information. Kognitiv einfache bzw. undifferenzierte Urteiler neigen im allgemeinen dazu, simultan dargebotene widersprüchliche Information zu einem konsistenten, wenig dimensional differenzierten Urteil zu verknüpfen. Bei sukzessiv dargebotener widersprüchlicher Information ändern sie ihr Urteil stark von einer Informationsdarbietung zur anderen, während kognitiv komplexe Urteiler relativ schnell eine ausgewogene Stabilität im Urteil erreichen.

Die große Anzahl von Untersuchungen zur Frage, welchen Einfluß die kognitive Komplexität auf das Urteil unter verschiedenen Bedingungen, wie variierendes Stimulusmaterial, Wiederholungen des Urteils usw. hat, trägt jedoch nicht ohne weiteres zur Strukturierung des Problemkreises bei, da die unterschiedlichen angewendeten Methoden zur Erfassung der kognitiven Komplexität nur wenig kovariieren (VANNOY, 1965), was aufgrund der Überlegungen von CAMPBELL und FISKE (1959) und der sich anschließenden Untersuchungen zu erwarten war.

Ein weiteres Problem, das sich bei der Beurteilung von Personen anhand von Skalen ergibt, ist das des sogenannten Halo-Effekts und des logischen Fehlers. Der Halo-Effekt tritt nach HOFSTÄTTER (1963; im Anschluß an THORNDIKE, 1918) vor allem dann auf, wenn Personen aufgrund minimaler Information beurteilt werden. Er schlägt sich in einer Erhöhung der Skaleninterkorrelationen nieder. Es wird angenommen, daß die Interkorrelationserhöhung insbesondere durch bewertende Beurteilung anhand der Skalen bedingt sei.

Skaleninterkorrelationen, die eher die Ansicht der Beurteiler über den Zusammenhang der Skalen als den tatsächlichen Zusammenhang widerspiegeln, bezeichnet man im Anschluß an NEWCOMB (1931) als logischen Fehler. Auch er soll sich insbesondere bei wenig Information in einer Erhöhung der Skaleninterkorrelationen äußern, womit impliziert ist, daß die Ansicht über Eigenschaftszusammenhänge - zumindest bei geringer Information - undifferenzierter ist als es den realen Zusammenhängen entspricht. Faktorenanalytisch ist in beiden Fällen eine Erhöhung der gemeinsamen Varianz, eine Varianzverstärkung des ersten Faktors und in extremen Fällen eine Verminderung der Anzahl gemeinsamer Faktoren zu erwarten.

Aus dem Konzept des Halo-Effekts sowie dem des logischen Fehlers resultiert, daß sich mit zunehmender Information die Höhe der Skaleninterkorrelationen vermindert, was sich entsprechend faktorenanalytisch in einer Erhöhung der spezifischen Varianz, der Anzahl gemeinsamer Faktoren und einer Verminderung der Varianzstärke des ersten Faktors auswirkt.

Erwartungsgemäß wird in Untersuchungen (KOLTUV, 1962; COHEN, 1969) bei zunehmender Information eine Verminderung der Höhe der Skaleninterkorrelationen nachgewiesen. Das Konzept des Halo-Effekts wird jedoch insofern in Frage gestellt, als auch nach Auspartialisierung bewertender Stellungnahmen (Sympathieurteil) eine überzufällige Anzahl signifikanter Korrelationen festgestellt wurde (KOLTUV, 1962). Damit scheint das Konzept des logischen Fehlers angemessener, jedoch führt das Wort "Fehler" in die Irre, da die Erfassung der tatsächlichen Beziehungen außerordentlich schwierig ist, so daß ein Vergleich "wahrer" und "falscher" Beziehungen nicht ohne weiteres vorgenommen werden kann. Es erscheint angebrachter, die aufgrund von Beurteilungen ermittelten Beziehungen im Sinne der impliziten Persönlichkeitstheorie als selbständiges Untersuchungsobjekt zu betrachten (vgl. auch COHEN, 1969). Die entsprechenden Schlußfolgerungen waren schon bei der Streitfrage, ob es sich um reale oder semantische Beziehungen handelt, gezogen worden. Und tatsächlich überschneidet sich die Fragestellung beider Problemkreise.

Wertende Stellungnahmen bleiben jedoch, insbesondere bei der Beurteilung mit Hilfe des semantischen Differentials, ein Problem. OSGOOD und Mitarbeiter (1957) haben aufgedeckt, daß bei der Beurteilung von Personen, aber auch von irgendwelchen Konzepten anhand des semantischen Differentials bewertende Stellungnahmen im Sinne der Beurteilung "gut vs. schlecht" usw. eine wichtige Beurteilungsdimension darstellen. Entsprechend haben ANDERSON und Mitarbeiter sowie eine große Anzahl weiterer Forscher den Effekt der Kombination simultan bzw. sukzessiv dargebotenen konsistenten bzw. inkonsistenten Stimulusmaterials auf der Bewertungsdimension "sympathisch vs. unsympathisch" untersucht (zusammenfassend ANDERSON, 1971). PEABODY (1967) stellt die Bedeutung der Bewertung bei der Personbeurteilung in Frage. Er zeigte auf, daß zwischen beschreibenden und bewertenden Aspekten der zur Beurteilung benutzten Eigenschaftswörter unterschieden werden muß. So gehören z.B. die Adjektive "ernst" und "grimmig" zum selben Pol der Beschreibungsdimension, ihre Bewertung ist jedoch gegensätzlich. Umgekehrt liegen die Adjektive "frivol" und "grimmig" auf den beiden Polen der Beschreibungsdimension, ihre Bewertung ist jedoch gleich, nämlich in beiden Fällen negativ.

In seiner Analyse versuchte er nachzuweisen, daß beschreibende Aspekte bei der Personbeurteilung wichtiger sind als bewertende. Dieses Ergebnis ist - auch aufgrund methodischer Überlegungen - von ROSENBERG und OLSHAN (1970) bezweifelt worden. FELIPE (1970) und PEABODY (1970) belegten in weiteren Untersuchungen, daß sowohl Beschreibungs- als auch Bewertungsaspekte eine Rolle spielen, Bewertungsaspekte jedoch im allgemeinen im Vordergrund stehen.

B. FRAGESTELLUNG

In der bisherigen Einleitung wurden einige der als wichtig erachteten Aspekte psychiatrischen Urteilens - im Sinne des Abgebens einer Diagnose - und psychologischen Urteilens anhand von Skalen aufgeführt und Unterschiede zwischen psychiatrischem und psychologischem Urteilen kurz skizziert. Ein Ziel dieser Untersuchung ist die Analyse der von Psychiatern getroffenen psychiatrischen und psychologischen Urteile über Patienten. Besonders interessiert dabei die Wirkung von zunehmender und gleichzeitig psychiatrisch wichtiger werdender Information aus Explorationen sowie der Effekt verschiedener zu beurteilender Patientengruppen auf das Urteil. Weiterhin sind interindividuelle Differenzen zwischen den Urteilern als dritte Bedingung von Interesse. Die Untersuchung verfolgt weniger theoretische Absichten, sondern versucht eine möglichst praxisnahe Analyse des psychiatrischen und des von Psychiatern ebenfalls sehr häufig vorgenommenen psychologischen Urteilens. Daher ist es nicht möglich, die beiden genannten Bedingungen, zunehmende Information und psychiatrisch relevanter werdende Information, zu trennen, um jeweils ihren Effekt auf das psychiatrische und psychologische Urteil zu überprüfen. Eine Koppelung der beiden Bedingungen liegt nämlich in der Praxis gewöhnlich immer vor, da - wie anfangs aufgezeigt - die Untersuchung eines unbekannten Patienten im allgemeinen mit der Kontaktaufnahme beginnt, um nach biographischer und sozialer Anamnese mit der spezifischen Exploration zur Erkrankung zu enden. Außerdem bringt ein Untersuchungsansatz, in dem beide Bedingungen getrennt werden, ganz erhebliche technische Schwierigkeiten mit sich. Korrekterweise können daher im folgenden aber auch nur Aussagen über den gleichzeitigen Einfluß beider Bedingungen auf das psychiatrische und psychologische Urteil gemacht werden. Ein Verlaufsvergleich des psychologischen und psychiatrischen Urteils soll Aufschluß über unterschiedliche kognitive Prozesse beim psychologischen und psychiatrischen Urteilen liefern. Dieser Vergleich muß im psychologischen Urteil die Patiententypen betreffen, da im psychiatrischen Urteil nur Patientenklassen erfaßt werden. Die Veränderung des psychologischen Urteils oder klarer formuliert: die Veränderung der

psychologischen Sicht der Patienten schlägt sich in einer Veränderung der Patiententypen nieder ; die Änderung der psychiatrischen Sicht der Patienten dagegen in einem Wechsel der Diagnosen.

Psychologisches und psychiatrisches Urteilen kann im Anschluß an die bisherigen Ausführungen als Kategorisierungsvorgänge beschrieben werden. Der Unterschied zwischen beiden Urteilsformen dürfte vorrangig im Ausmaß der Explizitheit des Urteilsprozesses bestehen. Psychiatrisches Diagnostizieren erfolgt im allgemeinen auf einer kognitiv relativ hohen Bewußtseinsstufe, auf der die vom Patienten gelieferte Information abstrahierend zu Symptomen zusammengefaßt und diese ermittelten Symptome mit den Diagnoseklassen definierenden verglichen werden; bei Übereinstimmung erfolgt die Kategorisierung in eine Diagnoseklasse. Daher ist zu erwarten, daß das psychiatrische Urteil in Abhängigkeit von zunehmender, psychiatrisch relevanter werdender Information starken Veränderungen unterliegt. Desgleichen dürfte für ein subjektiv sicheres und valides psychiatrisches Urteil viel und insbesondere psychiatrisch wichtige Information notwendig sein. Dagegen kann beim psychologischen Urteil erwartet werden, daß es in geringerem Grade von Informationsausmaß und -inhalt abhängig ist und sich entsprechend relativ schnell stabilisiert, da der zur Beurteilung anhand der Skalen führende Prozess auf einer kognitiv niederen Bewußtseinsstufe erfolgt, auf der neu hinzukommende, psychiatrisch relevanter werdende Information nicht mehr ausreichend korrekt verarbeitet wird. Diese Vermutung, das psychologische Urteil erhärte sich schnell, wurde mehrmals, innerhalb des klinischen Bereichs insbesondere durch MEEHL (1960) bestätigt. MEEHL wies nach, daß sich das Urteil von Therapeuten mit Hilfe Q-Sorts etwa zwischen der zweiten und vierten Therapiestunde verfestigt, d.h. die Q-Korrelationen waren ab zweiter bis vierter Stunde stabil.

Ein weiteres Ziel der Untersuchung ist die Analyse des Verhaltens explorierender Ärzte. Dabei interessiert insbesondere, ob sich das Verhalten in Abhängigkeit von den beiden Bedingungen Befragungsthema und Diagnosegruppe der explorierten Patienten ändert. Da ein entsprechendes, angemessenes Instrument zur Messung explorativen Verhaltens fehlt, wird es über das Urteil von Psychiatern anhand von Skalen erfaßt. Damit wird zwar die Objektivität von Aussagen über exploratives Verhalten eingeschränkt, jedoch können diese psychologischen Urteile wiederum unter Urteilsgesichtspunkten analysiert und diskutiert werden. Erwartet wird, daß sich insbesondere Unterschiede zwischen den explorativen Verhaltensweisen bei Kontaktaufnahme, biographischer und sozialer Anamnese

einerseits und bei spezifischer Exploration zur Erkrankung andererseits ergeben. Bezüglich des Einflusses der Patientengruppen bestehen keine konkreten Vermutungen. In allen nun folgenden Punkten exakte Hypothesen aus der Literatur stringent abzuleiten, ist nicht möglich, da insbesondere ein Verlaufsvergleich zwischen psychologischem und psychiatrischem Urteil in der Literatur fehlt. Es soll jedoch versucht werden, aufgrund der kurz dargestellten Ergebnisse aus der Literatur sowie der ausgeführten Überlegungen Hypothesen zu formulieren, wobei die das psychologische und psychiatrische Urteil beeinflussenden Bedingungen, Informationsausmaß und -inhalt, verschiedene zu beurteilende Patientengruppen und verschiedene beurteilende Psychiater, die Erwartungen bestimmen. Den Hypothesen folgen jeweils kurze Begründungen.

Allgemeine Gesichtspunkte: Durchschnittliches Urteil der Psychiater

- 1) Die psychologischen Beurteilungsdimensionen als kognitiver Beurteilungsraum sind inhaltlich gegenüber Informationsausmaß und -inhalt relativ invariant. Dieser Vermutung entsprechen die Überlegungen von KELLY (1955) und insbesondere das Modell von SARBIN et al. (1960). Darüber hinaus ist aufgrund des geringen Einflusses der Beurteilten auf die Urteilsstruktur (z.B. NORMAN und GOLDBERG, 1966; PASSINI und NORMAN, 1966; COHEN, 1969) zu erwarten, daß zunehmende, inhaltlich verschiedenartige Information über dieselben Personen einen eher noch geringeren Effekt ausübt, wenn - wie in dieser Untersuchung - die Informationsart (gewonnen durch Exploration) konstant bleibt.
- 2) Die aufgrund des psychologischen Urteils ermittelten, innerhalb des kognitiven Beurteilungsraums lokalisierten Patiententypen, die die psychologische Sicht der Patienten widerspiegeln, sind inhaltlich weniger resistent gegenüber Informationsausmaß und -inhalt als die Beurteilungsdimensionen, stabilisieren sich aber relativ schnell. Dieser Hypothese liegt wiederum das Modell von SARBIN et al. (1960) zugrunde. Patienten werden bei zunehmender, psychiatrisch wichtiger werdender Information psychologisch anders gesehen und entsprechend beurteilt. Die verhältnismäßig schnelle Stabilisierung läßt sich u.a. aus der Untersuchung von MEEHL (1960) sowie der ausgeführten Überlegung ableiten, psychologisches Urteilen erfolge auf einer - im Verhältnis zum psychiatrischen Urteil - kognitiv niederen Bewußtseinsstufe.
- 3) Es erfolgt eine Ausdifferenzierung der Urteilsstrukturen und insbesondere der Patiententypen in Abhängigkeit von zunehmender, psychiatrisch relevanter werdender Information.

Die vermutete Ausdifferenzierung der Urteilsstruktur steht zwar nicht unbedingt im Einklang mit dem Modell von SARBIN et al. (1960), entspricht aber den Annahmen zum Halo-Effekt und logischen Fehler und wird durch die entsprechenden Untersuchungen nahegelegt; in den Untersuchungen von KOLTUV (1962) und COHEN (1969) wurde bei zunehmender Information eine Verminderung der Höhe der Skaleninterkorrelationen nachgewiesen.

Die Erwartung, die Patiententypen differenzierten sich aus, basiert auf der Überlegung, daß zunehmende Information zu einer spezifischeren psychologischen Sicht der Patienten führt, d.h. es werden zunehmend feinere Unterschiede zwischen Patienten wahrgenommen.

4) Die aufgrund des psychologischen Urteils ermittelten Patiententypen stimmen nur wenig mit psychiatrischen Patientenklassen überein.

Diese Erwartung ist direkt aus der Literatur ableitbar (z.B. LORR et al., 1963). Die Unterschiede gehen auf formal unterschiedliche Klassen und - in Abhängigkeit davon - auf inhaltlich und formal unterschiedliche Beurteilungsskalen zurück. Darüber hinaus dürften - wie aufgezeigt - die kognitiven Beurteilungsvorgänge differieren.

5) Das psychiatrische Urteil erweist sich als relativ stark von Informationsausmaß und -inhalt abhängig, d.h. die psychiatrische Sicht der Patienten wechselt erheblich, so daß häufig Diagnoseveränderungen vorkommen. Entsprechend sind Wiederholungsreliabilitäten (Stabilität) gering, und die Validität steigt mit zunehmender und psychiatrisch wichtiger werdender Information deutlich an.

Diese Hypothese des starken Diagnosewechsels bzw. der geringen Stabilität folgt eindeutig aus Untersuchungen (z.B. ZUBIN, 1967). Darüber hinaus entspricht sie der Überlegung, daß das psychiatrische Urteil auf einer kognitiv hohen Bewußtseinsebene abläuft und neue Information verhältnismäßig exakt aufgenommen und verarbeitet wird. Da die neue Information zugleich zunehmend psychiatrisch wichtiger wird, resultiert das Anwachsen der Validität des psychiatrischen Urteils.

6) Aus Hypothese 5) und Hypothese 2) folgt, daß sich das psychologische Urteil schneller stabilisiert als das psychiatrische. Oder klarer formuliert: Die aus dem psychologischen Urteil ermittelbaren Patiententypen unterliegen inhaltlich relativ schnell einer Verfestigung. Dagegen ist die Anzahl von Diagnosewechseln als Anzeichen für die inhaltliche Labilität des psychiatrischen Urteils in Abhängigkeit von zunehmender, psychiatrisch

wichtiger werdender Information groß. Diese Erwartung entspricht insbesondere den bisherigen Ausführungen bezüglich des Unterschieds zwischen psychologischem und psychiatrischem Urteilsverlauf.

Der mögliche Einwand, der postulierte Verlaufsunterschied zwischen psychologischem und psychiatrischem Urteil resultiere aus der laut Untersuchungsansatz festgelegten Koppelung zwischen Zunahme und psychiatrischer Relevanz der Information, ist nicht stichhaltig. Zwar wird aufgrund dieser Koppelung eine besondere Labilität des psychiatrischen Urteils begünstigt, jedoch die Labilität des psychologischen Urteils nicht verhindert.

7) Patientengruppen als Stimulusbedingungen haben eindeutig Einfluß auf die Validität des psychiatrischen Urteils.

Diese Hypothese resultiert aus bisherigen, aufgeführten Untersuchungen (z.B. ZUBIN, 1967).

Differentielle Gesichtspunkte: Unterschiede zwischen urteilenden Psychiatern.

- 1) Die Beurteilungsstrukturen differieren interindividuell deutlich bezüglich Differenzierungsgrad und Inhalt.
- 2) Desgleichen differieren die Patiententypen interindividuell bezüglich Differenzierungsgrad und Inhalt.
- 3) Bei Stabilität und Validität des psychiatrischen Urteils treten eindeutige interindividuelle Unterschiede auf.

Ausführlicher Begründungen zu den drei Hypothesen bedarf es nicht. Interindividuelle Unterschiede wurden - wie in der Literaturübersicht aufgezeigt - sowohl bezüglich psychologischen als auch psychiatrischen Urteilens durch verschiedene Untersuchungen nachgewiesen.

Explorationsverhalten

- 1) Es werden zwei Explorationsstile erwartet. Der eine ist durch die Befragungsthemen Kontaktaufnahme, biographische und soziale Anamnese charakterisiert, der andere durch die Exploration zur Erkrankung. Diese Erwartung resultiert aus der eingangs ausgeführten Überlegung, daß bei der Exploration zur Erkrankung spezifische Symptome abgefragt wer-

den und somit die Exploration relativ straff verläuft. Dagegen dürfte der Befragungsstil insbesondere bei Kontaktaufnahme und sozialer Anamnese weniger gerichtet sein, d.h. die Befragung erfolgt eher im Sinne eines Gesprächs. Weiterhin bedarf die Information aus diesen Befragungsbedingungen - im Verhältnis zur Information aus der Exploration zur Erkrankung - im stärkeren Maße der Interpretation. Dieser subjektiv empfundene Unterschied im Ausmaß der Interpretation dürfte ebenfalls im Urteil der Psychiater zur erwarteten Unterscheidung zwischen beiden Explorationsstilen beitragen. Hinzu kommt, daß die Exploration zur Erkrankung den spezifisch psychiatrischen Explorationsstil darstellt; allein diese Tatsache dürfte die erwartete Unterscheidung im Urteil bedingen.

Welcher Explorationsstil Anwendung bei bestimmten Patientengruppen finden wird, läßt sich aufgrund der bisherigen Ausführungen nicht klar genug vorhersagen. Möglich ist, daß der durch die Exploration zur Erkrankung gekennzeichnete, spezifisch psychiatrische Explorationsstil entsprechend auch bei typisch psychiatrischen Patienten benutzt wird.

2) Die laut Urteil der Ärzte ermittelbaren Explorationsstile und die zur Beurteilung herangezogenen Beurteilungsdimensionen differieren interindividuell bezüglich Differenzierungsgrad und Inhalt.

Diese Hypothese folgt wiederum aus der genannten Literatur, laut der im psychologischen Urteil beträchtliche interindividuelle Differenzen auftreten.

C. UNTERSUCHUNG

1) Versuchsplan und Versuchsdurchführung

5 Psychiater explorierten insgesamt 15 Patienten. Pro Woche fand die Exploration eines Patienten statt. Das Patientengut setzte sich aus 5 neurotischen (N), 5 schizophrenen (S) und 5 cyclothymen (Z) Patienten zusammen. Die drei Diagnosegruppen fungieren als das Urteil über die Patienten sowie das Explorationsverhalten der Ärzte beeinflussende Bedingungen. Gleichzeitig wird mit der Auswahl der drei wichtigsten psychiatrischen Diagnosegruppen der Forderung Brunswiks entsprochen, ein möglichst repräsentatives Stimulusmaterial zu verwenden.

Die Patienten selbst sowie deren Diagnose waren den Psychiatern unbekannt.

Im folgenden werden die Gesamtuntersuchung eines Patienten als Exploration und einzelne Abschnitte innerhalb der Exploration als Explorations-Situationen bezeichnet.

Jeder Patient wurde nacheinander von jedem Psychiater in der Gegenwart der anderen Psychiater exploriert. Somit zerfiel die Exploration eines jeden Patienten in 5 aufeinanderfolgende Abschnitte (Situationen).

Die 5 Situationen waren durch folgende Explorationsthemen definiert:

- Situation I : Kontaktaufnahme; hier wurden Fragen gestellt wie: wie heißen Sie?, wo wohnen Sie?, wie sieht es dort aus? usw.
- Situation II : biographische Anamnese (Fragen nach körperlichen und insbesondere hirnorganischen Erkrankungen)
- Situation III : soziale Anamnese (Fragen nach sozialem und beruflichen Werdegang; soziale Beziehungen zu Eltern, Ehepartner usw.)
- Situation IV : Exploration zur Erkrankung
- Situation V : Exploration zur Erkrankung

Die Ärzte waren gehalten, während der Exploration in einer der Situationen nur zum entsprechenden Thema Fragen zu stellen und die Explorationsdauer in Situation I von 1 min., in Situation II von 3 min. und in den Situationen III bis V von jeweils 10 min. nicht zu unterschreiten.

Die durchschnittliche gesamte Explorationsdauer belief sich auf 37,7 min.. Pro Explorations-Situation betrug die durchschnittliche Dauer und die Streubreite der Explorationszeiten:

- Situation I : 1,5 min.; 0,8 - 2,9 min.
- Situation II : 3,9 min.; 1,1 - 12,3 min.
- Situation III : 8,5 min.; 2,7 - 16,2 min.
- Situation IV : 12,4 min.; 5,8 - 20,5 min.
- Situation V : 11,0 min.; 5,9 - 18,6 min.

Die jeweils explorierenden Ärzte verteilten sich pro Patientengruppe entsprechend folgendem Lateinischen Quadrat auf die 5 durch die Explorationsthemen definierten Situationen:

Situationen			I	II	III	IV	V
Patienten	1	Arzt	1	2	3	4	5
	2		5	3	2	1	4
Neurotiker	3		2	5	4	3	1
	4		4	1	5	2	3
	5		3	4	1	5	2
Schizophrene			entsprechend der Abfolge wie bei Neurotiker				
Cyclothyme			entsprechend der Abfolge wie bei Schizophrene				

Die Reihenfolge der Patienten, die sich der Exploration zu unterziehen hatten, ergab sich aus dem Angebot der Klinik (Neuropsychiatrische Klinik der Universität Giessen). Jedoch wurde versucht, Patienten unterschiedlicher Diagnosegruppen aufeinander folgen zu lassen.

Der dargestellte Untersuchungsansatz ist insofern realitätsnah, als die gemeinsame Exploration eines Patienten durch mehrere Psychiater eine in der Klinik durchaus übliche Explorationssituation ist. So werden z.B. fachlich besonders interessante Patienten in Konferenzen vorgestellt und gemeinsam exploriert. Auch die Explorationsthemen sowie deren Abfolge entsprechen der Praxis. Etwas unrealistisch ist jedoch die dem Lateinischen Quadrat entsprechende Verteilung der Ärzte auf die Explorations-Situationen und Patienten einer Patientengruppe, da die Ärzte damit gezwungen werden, nur Fragen zum jeweiligen Explorationsthema zu stellen. Damit ist – pro exploriertem Patient dem einzelnen Arzt die Möglichkeit genommen, ihn interessierende oder unklar gebliebene Fragen in denjenigen Situationen zu stellen, in denen seine Kollegen explorieren. Dieser Untersuchungsansatz ist jedoch notwendig, um Positions- und Abfolgeeffekte der explorierenden Ärzte zu vereinheitlichen. Es handelt sich um einen Plan, in dem auch die gegenseitige Beeinflussung der explorierenden Ärzte durch ihre jeweiligen Explorationen ausbalanciert ist.

Von den Psychiatern wurden folgende Beurteilungen der Patienten abgegeben:

a) Jeder Patient wurde anhand eines zur Beurteilung von Patienten entwickelten Polaritäten-Profiles (Sematisches Differential) nach den Explorations-Situationen I, III und V beurteilt. Somit wird der Einfluß zunehmender und psychiatrisch relevanter werdender

Information erfaßbar. Weiterhin wurde anhand einer Skala angegeben, in welchem Ausmaß die Entscheidung beim Anstreichen der Skalen leicht bzw. schwer fiel. (Das Polaritäten-Profil eignet sich sehr gut zur Beurteilung von Personen; siehe die Zusammenfassung von WARR und KNAPPER, 1968). Die Psychiater waren laut Instruktion gehalten, in ihrem Urteil ein Bild von den Patienten zu geben, das sowohl aktuelle als auch habituelle Verhaltens- und Persönlichkeitskomponenten umfaßt. Die kombinierte Aufgabe bereitete laut Angaben der Ärzte häufig Schwierigkeiten. Die Zusammenfassung aktueller und habitueller Komponenten war notwendig, um ein etwa gleiches Generalisationsniveau des psychologischen und psychiatrischen Urteils zu gewährleisten.

b) Jeweils nach Beurteilung anhand des Profils hatten die Ärzte die Patienten psychiatrisch zu diagnostizieren und das Ausmaß der subjektiven Sicherheit bezüglich der getroffenen Diagnosen zu beurteilen. Jeder Patient erhielt also 3 mal 5 psychologische und psychiatrische Beurteilungen.

Der Entscheidung, bei 5 Explorations-Situationen nur 3 Beurteilungen vornehmen zu lassen, lag die Absicht zugrunde, sehr starke Interferenzen zwischen den jeweiligen Beurteilungen zu vermeiden. Außerdem ist aus Untersuchungen bekannt, daß sich bei jeweiliger Beurteilung nach neuer Information ein sogenannter Recency-Effekt, dagegen bei nur einmaliger Beurteilung am Ende der sukzessiv dargebotenen Information ein sogenannter Primacy-Effekt auswirkt, wenn die sukzessive Information widersprüchlich ist. Mit dieser dreimaligen Beurteilung, die gewissermaßen eine Mittelstellung bezüglich der genannten Untersuchungsansätze einnimmt, wurde die Ausbalancierung möglicher Effekte solcher untersuchungstechnischen Ansätze angezielt.

Die Beurteilung der Patienten nach der ersten Explorations-Situation, also nach der Kontaktaufnahme, gibt den sogenannten ersten Eindruck der Ärzte wieder. Verbal vermittelte psychiatrisch wichtige Informationsinhalte stehen den Psychiatern bei dieser ersten Beurteilung noch nicht zur Verfügung. Die psychiatrische sowie psychologische Beurteilung der Patienten orientiert sich hier an motorischen und mimischen Verhaltensweisen und an der Form verbaler Reaktionen. Damit unterscheidet sich diese Beurteilung von den folgenden insbesondere durch minimale Informationsinhalte.

Bei der zweiten Beurteilung nach der Explorations-Situation III verfügen die Psychiater über zusätzliche Information aus biographischer und sozialer Anamnese. Die Zusammenfassung der Information aus diesen beiden Explorations-Situationen für das zweite Urteil ist insofern gerechtfertigt, als die Information aus der biographischen Anamnese zur Be-

urteilung der den drei Diagnosegruppen angehörenden Patienten eher nebensächlich ist. Die biographische Anamnese mußte jedoch in die Exploration aufgenommen werden, um eine vollständige Exploration zu erreichen. Das Weglassen dieser Explorations-Situation hätte möglicherweise eine spezifische Erwartung der Ärzte bezüglich der möglichen Diagnosen zur Folge gehabt. Außerdem ist das explorative Verhalten der Ärzte in dieser Explorations-Situation von Interesse.

Ausschlaggebend für das zweite psychologische und psychiatrische Urteil dürfte jedoch die Information aus der Explorations-Situation III, die Information aus der sozialen Anamnese, sein.

Das dritte und damit letzte Urteil erfolgt nach der gesamten Information über die Patienten und stützt sich damit auch auf die Information aus der Exploration zur Erkrankung. Diese Information ist für eine psychiatrische Diagnose von besonderer Wichtigkeit.

Die Beurteilung des Explorations-Verhaltens der Psychiater verlief folgendermaßen: Nach jeder der Explorations-Situationen wurde der Arzt, der die Exploration durchgeführt hatte, anhand eines eigens entwickelten Polaritäten-Profiles eingeschätzt (1 Selbst- und 4 Fremdbeurteilungen). Somit wurde jeder Arzt in jeder der 5 Explorations-Situationen dreimal beurteilt, nämlich nach Exploration eines Neurotikers, eines Schizophrenen und eines Cyclothymen. Insgesamt erhielt also jeder Arzt 15 Selbst- und 4 mal 15 Fremdbeurteilungen. Umgekehrt gab jeder Arzt 15 Selbst- und 4 mal 15 Fremdbeurteilungen ab.

Die Einschätzung der in den Explorations-Situationen I, III und V explorierenden Ärzte erfolgte jeweils nach Beurteilung des jeweiligen Patienten.

Nach Abschluß der Untersuchung, die sich über ein Wintersemester erstreckte, gaben die 5 Ärzte aus ihrer Vorstellung heraus Urteile über folgende generalisierten, typischen Patienten anhand des entsprechenden Polaritäten-Profiles ab:

1) Neurotiker mit depressiver Symptomatik, 2) Neurotiker mit hypochondrischer Symptomatik, 3) Neurotiker mit hysterischer Symptomatik, 4) Schizophrener mit ersten Schub, 5) Defekt-Schizophrener, 6) Hebephrener, 7) Cyclothymen mit manischer Phase, 8) endogen Depressiver und 9) Patient mit hirnanorganischem Psychosyndrom.

Außerdem beurteilten die Ärzte - ebenfalls aus der Vorstellung heraus - sich selbst und ihre 4 Kollegen bezüglich ihres durchschnittlichen, typischen Explorations-Verhaltens.

2) Beurteilungsinstrument: Polaritäten-Profil

Bei der Zusammenstellung der Adjektivpaare für das Polaritäten-Profil zur Beurteilung von Patienten fanden hauptsächlich die Adjektive aus sogenannten Markierungsbögen, die in verschiedenen Psychiatrischen Kliniken benutzt werden, Verwendung. Z.T. orientierte sich die Auswahl der Adjektivpaare auch an Adjektiven, die Psychiater in Krankengeschichten verwenden. Darüber hinaus wurden verschiedene Eigenschaftswörterlisten durchgesehen. Bei der Auswahl der Adjektive wurde auf Repräsentativität in Bezug auf relevante Eigenschaftsdimensionen von Patienten geachtet.

Die in dieser Untersuchung verwendeten Adjektive kennzeichnen relativ abstrakte Beschreibungsdimensionen. Zur Beurteilung von Patienten anhand dieser Eigenschaften aufgrund der Information aus der Exploration der Patienten bedarf es somit interpretativer Zwischenschritte. Da mit dieser Untersuchung u.a. die Analyse des praktischen psychologischen Urteilens beabsichtigt ist und solche Adjektive tatsächlich in der Praxis benutzt werden (siehe Krankengeschichten, Markierungsbögen), erscheint die Auswahl dieser Adjektivpaare gerechtfertigt. Es muß jedoch darauf hingewiesen werden, daß solche Eigenschaften den Einfluß von Interpretations- und Beurteilungstendenzen im Verhältnis zu verhaltensnahen Eigenschaften besonders begünstigen.

In einer Voruntersuchung wurde das Polaritäten-Profil von 46 Psychiatern verschiedener deutscher psychiatrischer Universitäts-Kliniken angewendet. Die Daten wurden faktoranalytisch ausgewertet und einige Skalen (Adjektivpaare) nach dimensional Gesichtspunkten verändert. Diese revidierte Form des Profils mit 21 Skalen (siehe Anhang I) wurde in der Hauptuntersuchung eingesetzt. Die Hauptuntersuchung dient somit auch der erneuten Überprüfung des Profils.

Bei der Zusammenstellung der Adjektivpaare zur Beurteilung des explorativen Verhaltens der Ärzte ergaben sich Schwierigkeiten, da aus der Literatur nur wenige entsprechende Adjektive bekannt sind. Die Auswahl erfolgte daher hauptsächlich nach Beobachtung und Befragung explorierender Psychiater. Um eine annähernde Äquivalenz des Beurteilungsvorganges beim Einstufen von Patienten und Psychiatern zu gewährleisten, wurden relativ abstrakte Adjektive zusammengestellt. Die Äquivalenz wurde angezielt, da das psychologische Urteil über Patienten und Psychiater verglichen werden soll. Dieses Profil durchlief dieselbe Entwicklung wie das Profil zur Beurteilung von Patienten. Die in der Hauptuntersuchung eingesetzte Form des Profils umfaßt 21 Skalen (siehe Anhang II).

3) Versuchspersonen

a) Psychiater

An der Untersuchung nahmen 5 männliche Psychiater der Neuropsychiatrischen Universitätsklinik Giessen teil. Die durchschnittliche Dauer psychiatrischer Tätigkeit dieser Ärzte betrug 6,3 Jahre (im einzelnen 3, 3, 4,5, 6 und 15 Jahre). Im Durchschnitt hatten die Ärzte an 2,6 Nervenkliniken psychiatrisch gearbeitet (im einzelnen 2, 2, 2, 2 und 5 Nervenkliniken).

b) Patienten

Von den 15 Patienten waren 13 männlichen und 2 weiblichen Geschlechts. Die 13 männlichen Patienten waren - mit einer Ausnahme - stationär aufgenommene Patienten der Neuropsychiatrischen Universitätsklinik Giessen. Ihre Diagnose als eine der Untersuchungsbedingungen wurde durch die sie behandelnden Ärzte erstellt. Die beiden weiblichen Patienten waren Insassen des Psychiatrischen Landeskrankenhauses Giessen. Beide litten - laut Diagnose des behandelnden Arztes - an einer endogenen Depression. Diese beiden Patientinnen mußten in die Untersuchung aufgenommen werden, weil sich zur Zeit der Untersuchung männliche Patienten mit einer Cyclothymie in keinem der beiden Krankenhäuser befanden. Andernfalls wäre die Untersuchung durch die Wintersemesterferien unterbrochen worden.

4) Methode

Die Daten, die bei der psychologischen Beurteilung der Patienten gewonnen wurden, sind in viererlei Hinsicht klassifizierbar. Die vier Aspekte sind:

a) 21 Skalen des Patienten-Polaritäten-Profiles

b) 15 beurteilte Patienten

c) 5 beurteilende Ärzte

d) dreimalige Wiederholung der Beurteilungen, im folgenden als Urteil benannt.

Unter "Allgemeine Gesichtspunkte: Durchschnittliches Urteil der Ärzte" sind zunächst im Zusammenhang mit den Hypothesen die durchschnittlichen Skalenstrukturen sowie die durchschnittlichen Patiententypen pro Urteil von Interesse. Mit Hilfe der üblichen Faktorenanalyse (FA) können die beiden Faktorenmatrizen lediglich getrennt errechnet werden, wobei sich noch wegen der notwendigen Mittlung über die Patienten oder die Psychiater ein zusätzliches Problem ergibt. Eine Methode, die diese Probleme umgeht,

die gleichzeitige Ermittlung beider Faktorenmatrizen leistet und außerdem noch die jeweiligen Psychiatertypen zu erfassen gestattet, ist die 3 Modi-Faktorenanalyse (3 Modi-FA) von TUCKER (1964, 1966). Die sogenannten Modi sind die die Daten bestimmenden 3 Aspekte. Hier betreffen die Modi Skalen, beurteilte Patienten und urteilende Psychiater. Zusätzlich liefert diese Methode noch Aufschluß über den Zusammenhang zwischen Skalen-Dimensionen, Patiententypen und Psychiatertypen.

Der Vergleich von jeweils Skalen- und Patientenfaktorenstrukturen zwischen den Urteilen gestattet die Überprüfung der postulierten inhaltlichen Veränderung bzw. Stabilität der Skalen- und Patientenfaktorenstrukturen von Urteil 1 bis 3. Dieses Vorgehen trägt in gewisser Hinsicht einem der zentralen Kritikpunkte an der FA, nämlich der Rotation der Faktorenlösung, Rechnung (vgl. insbes. die Beiträge zum "Münchener Symposium über FA", 1967). An üblichen faktorenanalytischen Untersuchungen wird kritisiert, daß Faktorenmatrizen ohne Bezug auf irgendwelche Hypothesen rotiert und interpretiert werden. Folge dieser Vorgehensweise ist eine Vielfalt von unverbindlichen, zusammenhangslosen Untersuchungsergebnissen.

Schon BECHTHOLD (1961) stellte dieser erkundenden (exploratory) Anwendung der FA eine überprüfende (confirmatory) gegenüber. Bei der überprüfenden Anwendung wird die extrahierte Faktorenstruktur auf eine als Hypothese formulierte Faktorenstruktur rotiert und aufgrund des Ausmaßes an Übereinstimmung die Hypothese angenommen oder verworfen. Im Rahmen des Münchener Symposiums wird eine solche Hypothesenüberprüfende Anwendung der FA gefordert. Diese Vorgehensweise verlangt Hypothesengebäude vor Anwendung der FA, die übliche erkundende Anwendung der FA zielt dagegen auf Hypothesengewinnung ab. Da die faktorenanalytisch gewonnenen Ergebnisse jedoch in starkem Maße von Stichproben der Versuchspersonen sowie der verwendeten Variablen abhängig sind, kommt es bei der erkundenden Vorgehensweise zu der genannten Vielzahl von nicht integrierbaren Ergebnissen.

Die faktorenanalytische Methode von TUCKER wird in dieser Untersuchung gewissermaßen sowohl zu Erkundungs- als auch zu Überprüfungszwecken eingesetzt. So interessieren im Sinne der Erkundung die Interpretation der rotierten Skalen-, Patienten- und Psychiaterfaktoren. Im Sinne der Überprüfung erfolgt zwar kein Vergleich von extrahierten mit postulierten Faktorenstrukturen, doch wird mit Ähnlichkeitsvergleichen der Faktorenstrukturen aus den drei Urteilen der Hypothese über die Veränderung bzw. Stabilität des psychologischen Urteils nachgegangen.

Einen ähnlichen Dienst erweist die 3 Modi-FA bei der Analyse der durch die einzelnen Ärzte vorgenommenen Beurteilungen. Die pro Arzt die Daten bestimmenden 3 Modi sind dabei Skalen, beurteilte Patienten und die dreimalige Wiederholung des Urteils. Einerseits interessieren dabei Skalen-, Patienten- und Urteils-Dimensionen, andererseits werden mit Hilfe des Ähnlichkeitsvergleichs die jeweiligen individuellen Faktorenmatrizen auf interindividuelle Unterschiede hin überprüft.

Unter entsprechenden Gesichtspunkten erfolgt die Analyse des explorativen Verhaltens der Ärzte. Dabei sind die 3 Modi a) pro urteilendem Psychiater: Skalen des Arzt-Polaritäten-Profiles, beurteilte Psychiater und 15 malige Beurteilungen, b) pro beurteiltem Psychiater: Skalen, 15 malige Beurteilungen und Beurteiler.

Die 3 Modi-FA stellt eine Erweiterung der zu deskriptiven Faktoren führenden Hauptkomponentenmethode dar. Durch sie können - wie schon erwähnt - in dreierlei Hinsicht (Modi) klassifizierte Daten analysiert werden. Um die Ähnlichkeit zwischen der üblichen Hauptkomponentenmethode und dieser Methode aufzuzeigen, soll zunächst eine Umformung des Hauptkomponentenmodells vorgenommen werden.

Die fundamentale Bestimmungsgleichung des allgemeinen Faktorenmodells und somit auch die des Hauptkomponentenmodells ist

$$X = A B.$$

Sie besagt, daß die Datenmatrix X das Produkt aus der Faktorenmatrix A und der Faktorwertematrix B ist. Durch die Umformung kann folgende Gleichung erreicht werden:

$$X = A G B$$

Matrix G ist eine Kernmatrix, die die Interaktion zwischen den beiden Matrizen passender Dimensionalität A und B angibt (LEVIN, 1965). Diese Umformung basiert auf dem Theorem von ECKART und YOUNG (1936).

Bei der 3 Modi-FA wird von folgendem Modell ausgegangen, das in seiner allgemeinen Form in Summationsschreibweise dargestellt wird:

$$(1) x_{ijk} = \sum_{m,p,q} a_{im} b_{jp} c_{kq} g_{mpq}$$

Die Indizes bezeichnen:

i, i' : Index für beobachteten Probandenwert (bzw. Modus i)

j, j' : Index für beobachteten Skalenwert (bzw. Modus j)

k, k' : Index für beobachteten Beurteilerwert (bzw. Modus k)

Der mit einem Strich versehene Index bezeichnet ein Alternativelement.

m, m' : Index für Probandenfaktor (bzw. für Faktor des Modus i)

p, p' : Index für Skalenfaktor (bzw. für Faktor des Modus j)

q, q' : Index für Beurteilerfaktor (bzw. für Faktor des Modus k)

$\hat{x}_{ijk}$: approximierter beobachteter dreimodaler Wert; er entspricht dem beobachteten Wert x_{ijk} , wenn in die Analyse alle Eigenwerte größer Null eingehen.

a_{im} : Element der Matrix A; diese Koeffizienten beschreiben die beobachteten Elemente (i) in Termini der durch die Analyse abgeleiteten (m). Matrix A stellt die Matrix der Probandenfaktoren (Faktoren des Modus i) dar.

Entsprechend sind die Elemente b_{jp} und c_{kq} definiert.

g_{mpq} : Element der dreimodalen Kernmatrix G; diese Koeffizienten geben den Zusammenhang zwischen den Faktorenmatrizen A, B und C an.

Die fundamentale Gleichung des Modells kann für alle Werte als 2 Modi-Matrix dargestellt werden. Zur Darstellung wird das Kronecker-Produkt mitverwendet (BELLMAN, 1960)

Es ergeben sich drei Darstellungsweisen:

$$(2a) \quad {}_i \hat{X}_{(jk)} = {}_i A_m \quad G_{(pq)} \quad ({}_p B_i \times {}_q C_k)$$

$$(2b) \quad {}_i \hat{X}_{(ik)} = {}_i B_p \quad G_{(mq)} \quad ({}_m A_i \times {}_q C_k)$$

$$(2c) \quad {}_k \hat{X}_{(ij)} = {}_k C_q \quad G_{(mp)} \quad ({}_m A_i \times {}_p B_j)$$

Jede dieser drei Formen verwendet jeweils einen der drei beobachteten Modi als Reihen-Modus für die Matrix $\hat{X}$ und die beiden anderen beobachteten Modi als Kombinations-Modus für die 2 Modi-Matrix. Die Kernmatrix G ist der jeweiligen Matrix $\hat{X}$ angepaßt.

Ein Element der ersten der drei Matrizen stellt sich in Summationsschreibweise folgendermaßen dar:

$$(3) \quad \hat{x} = \sum_m a_{im} \quad \sum_p \sum_q g_{mpq} \quad (b_{jp} \quad c_{kq})$$

Die erste der drei Matrizen hat folgende Form, wobei festgelegt sei: $i = 5$, $j = 3$ und $k = 2$.

1j	2j	3j
1k 2k (1jk) (2jk)	1k 2k (3jk) (4jk)	1k 2k (5jk) (6jk)

1i

2i

3i

4i

5i

Die beiden anderen Matrizen können entsprechend veranschaulicht werden.

Es können folgende Produkt-Matrizen erstellt werden:

$$(4a) \quad \hat{M}_i = \hat{X}_{(jk)} \hat{X}_i$$

$$(4b) \quad \hat{P}_i = X_{ik} X_i$$

$$(4c) \quad \hat{Q}_k = \hat{X}_{(ij)} \hat{X}_k$$

Matrix $\hat{M}_i = \hat{X}_{(jk)} \hat{X}_i$ enthält als Elemente die Skalarprodukte aus der Multiplikation jeweils jeder Zeile der Matrix $\hat{X}_{(jk)}$ mit sich selbst und allen anderen Zeilen. Entsprechendes gilt für die Matrizen $\hat{P}_i$ und $\hat{Q}_k$.

Durch Substitution von (2) in (4) ergibt sich:

$$(5a) \quad \hat{M}_i = A_m G_{(pq)} (B_i \times C_k) (B_p \times C_q)_{(pq)} G_m A_i$$

$$(5b) \quad \hat{P}_i = B_p G_{(mq)} (A_i \times C_k) (A_m \times C_q)_{(mq)} G_p B_i$$

$$(5c) \quad \hat{Q}_k = C_q G_{(mp)} (A_i \times B_i) (A_m \times B_p)_{(mp)} G_q C_k$$

Es sei:

$$(6a) \quad M_m = G_{(pq)} (B_i \times C_k) (B_p \times C_q)_{(pq)} G_m$$

$$(6b) \quad P_p = G_{(mq)} (A_i \times C_k) (A_m \times C_q)_{(mq)} G_p$$

$$(6c) \quad q Q_q = q G_{(mp)} ({}_m A_i \times {}_p B_i) ({}_i A_m \times {}_i B_p) (mp) G_q$$

Dann gilt aus (5) und (6)

$$(7a) \quad {}_i \hat{M}_i = {}_i A_m M_m A_i$$

$$(7b) \quad {}_i \hat{P}_i = {}_i B_p P_p B_i$$

$$(7c) \quad {}_k \hat{Q}_k = {}_k C_q Q_q C_k$$

In diesen Gleichungen (7a - c) entsprechen die Matrizen ${}_i A_m$, ${}_i B_p$ und ${}_k C_q$

Faktorenmatrizen. Die Matrizen ${}_m M_m$, ${}_p P_p$ und ${}_q Q_q$ aus den Gleichungen (6a - c)

können als Kovarianzmatrizen zwischen den Faktoren interpretiert werden.

Sind die Spalten der Matrizen ${}_m A_i$, ${}_p B_i$ und ${}_q C_k$ orthonormal, dann gilt

$$(8a) \quad {}_m A_i A_m = {}_m I_m$$

$$(8b) \quad {}_p B_i B_p = {}_p I_p$$

$$(8c) \quad {}_q C_k C_q = {}_q I_q$$

Es handelt sich hier um auf den Betrag Eins normierte Faktoren.

Die Spalten des Kronecker-Produkts zweier solcher Matrizen sind dann selbst orthonormal. Daher werden die Gleichungen (6a - c) zu

$$(9a) \quad {}_m M_m = {}_m G_{(pq)} G_m$$

$$(9b) \quad {}_p P_p = {}_p G_{(mq)} G_p$$

$$(9c) \quad {}_q Q_q = {}_q G_{(mp)} G_q$$

Weiterhin gelten folgende Restriktionen

$$(10a) \quad \text{Rang } ({}_i \hat{X}_{(jk)}) = \text{Rang } ({}_i A_m) = \text{Rang } ({}_m G_{(pq)} ({}_p B_i \times {}_q C_k))$$

$$(10b) \quad \text{Rang } ({}_i \hat{X}_{(ik)}) = \text{Rang } ({}_i B_p) = \text{Rang } ({}_p G_{(mq)} ({}_m A_i \times {}_q C_k))$$

$$(10c) \quad \text{Rang } ({}_k \hat{X}_{(ij)}) = \text{Rang } ({}_k C_q) = \text{Rang } ({}_q G_{(mp)} ({}_m A_i \times {}_p B_j))$$

Unter Berücksichtigung dieser Restriktionen und der Tatsache, daß die Koeffizientenmatrizen ${}_i A_m$, ${}_i B_p$ und ${}_k C_q$ spaltenweise aus orthonormalen Vektoren bestehen, gilt

$$(11a) \text{ Spur } ({}_i \hat{M}_i) = \text{Spur } ({}_m M_m)$$

$$(11b) \text{ Spur } ({}_i \hat{P}_i) = \text{Spur } ({}_p P_p)$$

$$(11c) \text{ Spur } ({}_k \hat{Q}_k) = \text{Spur } ({}_q Q_q)$$

Die Kernmatrix kann nach Erstellung der Koeffizientenmatrizen über

$${}_m G_{(pq)} = {}_m A_i X_{(jk)} ({}_i B_p \times {}_k C_q)$$

oder

$${}_p G_{(mq)} = {}_p B_i X_{(ik)} ({}_i A_m \times {}_k C_q)$$

oder

$${}_q G_{(mp)} = {}_q C_k X_{(ij)} ({}_i A_m \times {}_i B_p)$$

errechnet werden (zur Ableitung siehe Gleichung (2)).

Gehen in die Analyse alle Eigenwerte größer Null ein, dann stimmen die rückgerechneten Werte genau mit den beobachteten überein. Da jedoch bei der Faktorisierung der Matrizen ${}_i M_i$, ${}_i P_i$ und ${}_k Q_k$ nicht alle Eigenwerte größer Null berücksichtigt, sondern entsprechend den üblichen Abbruchkriterien vorgegangen werden soll, ergeben sich folgende Schwierigkeiten:

Da jedes Element in der Kernmatrix durch jeden der 3 Modi bestimmt ist, werden über die Kernmatrix errechnete Eigenwerte bezüglich eines Modus (siehe Gleichung (9)) durch die (reduzierte) Faktorenanzahl der anderen Modi mitbestimmt. Daher entspricht die rückgerechnete Spur eines Modus nicht ganz der zu erwartenden. Weiterhin hat die Interaktion zwischen den Modi zur Folge, daß bei Weglassen von Elementen des einen Modus in der entsprechend reduzierten Kernmatrix die Produktmatrizen der anderen Modi nicht diagonal sind. Daher entsprechen Eigenwerte und Eigenvektoren nicht genau denen aus der Lösung der Matrix X. Aus diesem Grunde liefert diese Methode keine genaue Approximation an die beobachteten Daten nach dem Prinzip der kleinsten quadratischen Abweichungen.

Folgende Rechenschritte sind zur Ermittlung der Koeffizientenmatrizen und der Kernmatrix erforderlich:

1) Berechnet werden die Produktmatrizen ${}_i M_i$, ${}_i P_i$ und ${}_k Q_k$ aus der Matrix X der beobachteten Daten. Diese quadratischen Produktmatrizen führen auf das Eigenwertproblem hin.

2) Berechnet werden die Eigenwerte und Eigenvektoren dieser Produktmatrizen entsprechend der Hauptkomponentenmethode.

3) Es werden nur als signifikant erachtete Eigenwerte berücksichtigt, die bezüglich ihrer Größe in absteigender Rangfolge in den Matrizen ${}_m M_m$, ${}_p P_p$ und ${}_q Q_q$ stehen.

Die Frage nach der Anzahl der als signifikant zu betrachtenden Eigenwerte ist die übliche Frage nach einem Abbruchkriterium bei der Extraktion von Faktoren.

4) Die Koeffizientenmatrizen ${}_i A_m$, ${}_i B_p$ und ${}_k C_q$ werden gebildet. Die jeweils der Anzahl der berücksichtigten Eigenwerte entsprechenden Faktoren enthalten auf Einheitslänge normierte Vektoren (orthogonale Vektoren). Dadurch haben alle Faktoren dasselbe Gewicht. Das gilt auch für rotierte Faktoren.

5) Die Koeffizientenmatrizen werden varimaxrotiert. Auf die Kernmatrix hat die Rotation der Koeffizientenmatrizen natürlich Einfluß.

Versuche Tuckers, statt der Koeffizientenmatrizen die Kernmatrix auf Einfach-Struktur zu rotieren, waren nicht sehr erfolgreich.

6) Berechnet wird die Kernmatrix z.B. über die Gleichung

$${}_m G_{(pq)} = {}_m A_i X_{(jk)} ({}_i B_p \times {}_k C_q)$$

Es ist noch zu beachten, daß jeweils bei Erstellung der Matrizen ${}_i M_i$, ${}_i P_i$ und ${}_k Q_k$ die beiden anderen Modi in die Berechnung eingehen. Dieses Vorgehen impliziert gewissermaßen eine Reliabilitätssteigerung, was eine erhöhte Stabilität der errechneten Faktoren der jeweiligen Matrizen zur Folge hat. Dieser Effekt wird jedoch bei dem üblichen faktorenanalytischen Vorgehen zur klinischen Urteilsbildung auch nicht ausgeschaltet, da in diesem Fall vor der Berechnung über einen oder zwei Modi der Daten gemittelt wird, wodurch - im Gegensatz zu der Tucker'schen Analyse - Information verlorengeht.

Laut Tucker ist die 3 Modi-FA sowohl auf standardisierte als auch auf nichtstandardisierte Daten anwendbar, wobei die Standardisierung in Bezug auf einen Modus, 2 oder alle 3 Modi erfolgen kann. Es ist zu beachten, daß bei nichtstandardisierten Daten die Höhe der Produktsummen nicht nur von der Kovarianz zwischen jeweils zwei Variablen

bestimmt wird, sondern eindeutig auch vom Mittelwert und von der Streuung der Variablen. Der Effekt unterschiedlicher Mittelwerte und Streuungen ist nicht auspartialisierbar; Mittelwert, Streuung und Kovarianz bilden ein nicht trennbares Konglomerat, das sich entsprechend auch in den Faktoren niederschlägt. Die 3 Modi-FA nicht-standardisierter Daten widerspricht der Forderung CRONBACHs (1955), die 3 Komponenten impliziter Persönlichkeitstheorien Mittelwert, Streuung und Kovarianz voneinander getrennt in einer Untersuchung zu analysieren. Andererseits wird man die Frage nach Standardisierung der Daten nur unter Berücksichtigung des Ziels der jeweiligen Untersuchung beantworten können. Absicht dieser Untersuchung ist es, ein möglichst realitätsnahes Abbild psychologischen Urteilens zu liefern. Die Manipulation (Standardisierung) der Daten soll daher auf das unbedingt notwendige Maß beschränkt werden:

Bei der Auslese der Skalen aufgrund der Voruntersuchung wurde versucht, nur faktoriell eindeutige Skalen und Skalen mit einem Mittelwert nahe dem Wert 4 in das Polaritäten-Profil aufzunehmen. Die in der Hauptuntersuchung ermittelten Skalenmittelwerte wichen jedoch z.T. beträchtlich nach beiden Seiten vom Wert 4 ab und differierten somit erheblich. Diese Mittelwertsunterschiede dürften vorrangig Eigentümlichkeiten der Skalen selbst widerspiegeln; sie dürften nur zu einem geringen Ausmaß auf die anderen Modi zurückzuführen sein. Da in dieser Untersuchung außerdem nur die Struktur der zur Beurteilung benutzten Skalen interessiert, erschien der Effekt unterschiedlicher Skalenmittelwerte auf die durchzuführenden 3 Modi-FA nicht wünschenswert. Die Skalen wurden daher pro Analyse auf den gemeinsamen Mittelwert transformiert. In die Analyse gehen also Skalenabweichungswerte ein. Von einer Varianzhomogenisierung der z.T. recht varianzheterogenen Skalen wurde abgesehen, da die den einzelnen Skalen bei der Beurteilung zugemessene Gewichtung, als deren Ausdruck die Varianz interpretiert werden kann, im Sinne der Absicht der Untersuchung informativ ist. Mit der Vergrößerung der Skalenvarianz wächst die Wahrscheinlichkeit, daß die Skalen in den jeweiligen Faktoren hohe Ladungen erhalten. Eine hohe Ladung einer Skala in einem Faktor ist daher nicht nur ein Hinweis für die faktorielle Eindeutigkeit einer Skala, sondern gibt auch die Bedeutung der Skala für die beurteilenden Ärzte an.

Eine Standardisierung der Daten in Bezug auf die jeweiligen beiden weiteren Modi (beurteilte Patienten, beurteilende Ärzte oder beurteilte Patienten, Urteilswiederholungen oder beurteilende Ärzte, beurteilte Ärzte usw.) wurde aus folgenden Gründen

nicht vorgenommen:

Im Gegensatz zu Mittelwertunterschieden zwischen den Skalen enthalten Mittelwertunterschiede zwischen den Variablen der jeweiligen Modi Information. Die Ähnlichkeit zwischen den Variablen der jeweiligen Modi wird nämlich nicht nur durch die Konfiguration auf den Skalen, sondern ebenfalls durch Niveauunterschiede (Mittelwertunterschiede) bestimmt. Bei Profilvergleichen werden daher sowohl Korrelationen als auch Differenzmaße erstellt. Beide Komponenten gehen hier in die Berechnung ein. Desgleichen werden unterschiedliche Varianzen der Variablen der verschiedenen Modi als informativ angesehen. Z.B. sollen besonders auffällige und daher varianzstark eingeschätzte Patienten auch im Ergebnis der Analyse entsprechende Bedeutung haben (hohe Ladungen in den Patientenfaktoren).

D. ERGEBNISSE

I Beurteilung der Patienten

Als bevorzugte statistische Analyse-Methode des psychologischen Urteils wird die 3 Modi-FA verwendet. Die Ergebnisse sind geordnet in 1) Allgemeine Aspekte: Durchschnittliches Urteil der 5 Psychiater (in die 3 Modi-FA gehen alle 5 Ärzte ein) und 2) Differentielle Aspekte: Ärzte (pro Arzt wird eine 3 Modi-FA durchgeführt).

Innerhalb der allgemeinen Aspekte läßt sich sinnvollerweise weiterhin in a) Differenziertheit (Varianzanteil der Faktoren) b) Skalen, c) Patienten und d) Ärzte gliedern, da insbesondere bei b) und c) eine Weiterverarbeitung der durch die 3 Modi-FA gewonnenen Daten sowie derjenigen Werte erfolgen soll, die bei der aus der Vorstellung heraus getroffenen Beurteilungen typischer Patienten ermittelt wurden. Eine eindeutige Trennung dieser 3 Komponenten kann jedoch nicht vorgenommen werden, da sie Bestandteile einer 3 Modi-FA sind und in e) der Kernmatrix zusammengefaßt werden. Die Weiter- und Neuverarbeitung von Daten erfolgt jeweils innerhalb eines durch eine der drei Komponenten bestimmten Abschnitts.

Entsprechendes gilt für die Gliederung der Ergebnisse unter 2) Differentielle Aspekte: Ärzte.

Entsprechend wird das psychiatrische Urteil in "Allgemeine Aspekte" und "Differentielle

Aspekte" aufgegliedert. Um den Vergleich zwischen psychologischem und psychiatrischem Urteil zu erleichtern, werden zunächst die "Allgemeinen Aspekte" aus psychologischem und psychiatrischem Urteil dargestellt, anschließend folgen die jeweiligen "Differenziellen Aspekte".

Die Daten des psychiatrischen Urteils können unter "Allgemeine Aspekte" wegen der wahrscheinlichen Interaktion der Meßwertträger "beurteilende Psychiater, beurteilte Patienten" nicht statistisch analysiert werden. Eine solche statistische Analyse ist nur unter "Differenziellen Aspekten" möglich. Auf diese Problematik wird an entsprechender Stelle noch näher eingegangen werden.

In der Diskussion werden dann die unter "Differenziellen Aspekten" statistisch abgesicherten Ergebnisse auch z.T. zur Erläuterung der Ergebnisse unter "Allgemeine Aspekte" herangezogen. Das gilt insbesondere für die Ergebnisse unter dem Gesichtspunkt intraindividueller Unterschiede.

1. Allgemeine Aspekte

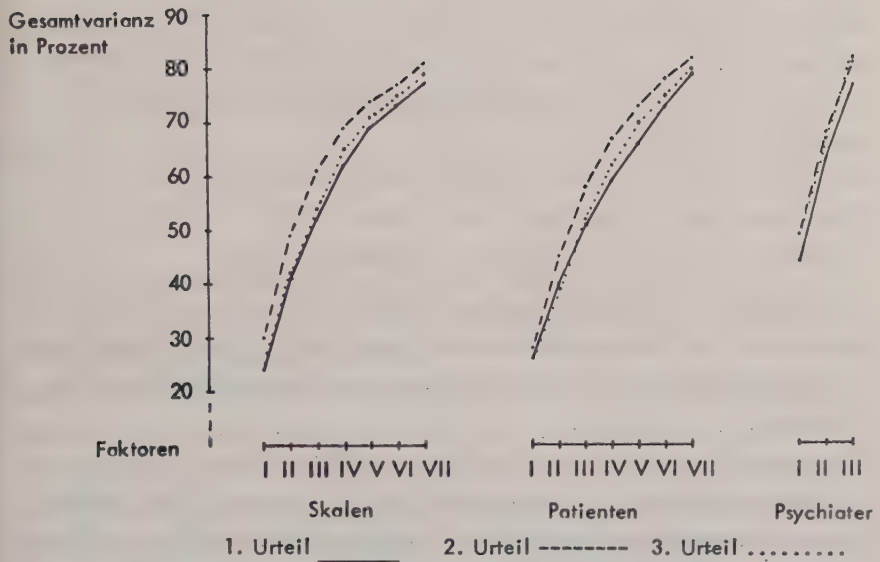
11. Psychologisches Urteil

Pro Urteil wurden die Daten jeweils einer 3 Modi-FA unterzogen. Die Modi sind Skalen, beurteilte Patienten und beurteilende Ärzte.

111. Differenziertheit des Urteils

Die Spur betrug im 1., 2. und 3. Urteil 1989,22 2167,75 und 2247,57. Die Spur ist für die drei Modi jeweils einer Analyse identisch. Sie ist Ausdruck für die Gesamtvarianz innerhalb eines jeden der drei Urteile. Danach wuchs die Gesamtvarianz von 1. zu 2. Urteil um 8,5%, von 2. zu 3. Urteil um 3,6% und von 1. zu 3. Urteil um 12,98% an. Die Erhöhung der Gesamtvarianz kann als zunehmende Sicherheit und Klarheit der durch die Ärzte getroffenen Beurteilungen interpretiert werden.

In der folgenden graphischen Darstellung 1 sind die Prozentanteile der pro Modus durch die unrotierten Faktoren erfaßten gesamten Varianz für das 1., 2. und 3. Urteil dargestellt. Die Anzahl der pro Modus als bedeutsam anzusehenden Faktoren und die Varianzstärke des 1. oder aller bedeutsamen Faktoren geben Aufschluß über das Ausmaß der Differenziertheit des Urteils unter dem jeweiligen Modus-Aspekt.



Graph. Darst. 1 : Kumulierte prozentuale Gesamtvarianzanteile der unrotierten Faktoren pro Modus und Urteil

Eine erste Betrachtung der graphischen Darstellung vermittelt, daß die pro Modus durch die Faktoren erfaßte Gesamtvarianz durchschnittlich von 1. zu 2. Urteil ansteigt, um dann von 2. zu 3. Urteil auf einen Wert zwischen 1. und 2. Urteil abzusinken. Dieser Sachverhalt trifft nicht durchgängig für die durch den ersten bzw. die beiden ersten Faktoren erfaßte Gesamtvarianz zu. Er wird jedoch bei der Anzahl derjenigen Faktoren pro Skalen- und Patienten-Modus deutlich, die aufgrund der Höhe ihrer Eigenwerte als bedeutsam zu betrachten sind. Die Anzahl der als bedeutsam erachteten Faktoren beträgt beim Skalen-Modus in allen drei Urteilen 4 Faktoren. Der kumulierte prozentuale Anteil der durch diese 4 Faktoren erfaßten Gesamtvarianz beläuft sich in allen drei Urteilen auf über 60% (siehe graphische Darstellung). Beim Patienten-Modus ist die Anzahl der als bedeutsam anzusehenden Faktoren pro Urteil nicht identisch. Im ersten Urteil treten eher nur drei, in den beiden folgenden Urteilen jeweils vier bedeutsame Faktoren auf. Dieses Ergebnis bestätigt die Hypothese, daß zunehmende Information zu differenzierteren Patiententypen im psycholo-

gischen Urteil der Psychiater führt. Aus Vergleichsgründen sollen die durch jeweils 4 Faktoren erfaßten Varianzanteile gegenübergestellt werden. Die genannte Varianzänderung tritt auch hier eindeutig auf; d.h. der durch 4 Faktoren erfaßte Varianzanteil wächst von 1. zu 2. Urteil an und fällt von 2. zu 3. Urteil auf einen Wert zwischen 1. und 2. Urteil. Ein entsprechendes Varianzverhältnis findet sich jedoch auch bei nur 3 Faktoren.

Bedeutsam ist in allen 3 Urteilen beim Arzt-Modus lediglich der 1. Faktor. Das beim Skalen- und Patienten-Modus beschriebene Varianzverhältnis ist bei einem Faktor nicht beobachtbar, jedoch bei zwei (siehe graphische Darstellung).

Die Tatsache, daß die Gesamtvarianz von 1. bis 3. Urteil ansteigt, während jeweils beim Skalen- und Patienten-Modus der durch 4 bzw. durch 3 Faktoren erfaßte Gesamtvarianzanteil von 1. bis 3. Urteil der genannten Schwankung unterliegt, bedarf der Erläuterung.

Ungeachtet differenzierterer statistischer Überlegungen kann zunächst erwartet werden, daß mit wachsender Gesamtvarianz von 1. bis 3. Urteil sich auch der faktoriell erfaßte Gesamtvarianzanteil erhöht. Dieser Sachverhalt tritt beim Skalen- und Patienten-Modus von 1. zu 2. Urteil ein und zwar sowohl bei der durch einen als auch durch 4 Faktoren extrahierten Gesamtvarianz. Die Hypothese kann jedoch beim Vergleich zwischen 2. und 3. Urteil nicht bestätigt werden. Die durch einen bzw. 4 Faktoren erfaßte Gesamtvarianz fällt bei beiden Modi. Beim Skalen-Modus ist der im 3. Urteil durch einen Faktor erfaßte Varianzanteil sogar noch etwas geringer als der entsprechende Varianzanteil im 1. Urteil.

Die Veränderung von 1. zu 2. Urteil entspricht der statistischen Erwartung. Diese Veränderung eindeutig als statistisches Artefakt zu interpretieren, ist nicht ganz gerechtfertigt, zumal der entsprechende, ebenfalls zu erwartende Effekt von 2. zu 3. Urteil nicht auftritt. Eine Redifferenzierung wird man jedoch im besten Fall nur beim Skalen-Modus vermuten können. Dieser möglichen Redifferenzierung folgt dann im 3. Urteil eine Differenzierung und wahrscheinlich auch eine Stabilisierung der Faktoren.

Dagegen ist beim Patienten-Modus eine fortschreitende Differenzierung zu beobachten, da sich im 2. Urteil trotz erhöhter extrahierter Gesamtvarianz der 4. Patienten-Faktor herauszukristallisieren beginnt. Diese Differenzierung setzt sich auch im 3. Urteil fort, und vermutlich erfolgt auch eine Stabilisierung der Faktoren.

Im Gegensatz zu den Ergebnissen aus Skalen- und Patienten-Modus finden sich beim Arzt-Modus keine Hinweise auf eine Ausdifferenzierung. Der Varianzanteil des 1. Faktors wächst von 1. zu 3. Urteil, dagegen fällt der Anteil des 2. Faktors im 3. Urteil. Daher erscheint es eher sinnvoll, eine Redifferenzierung anzunehmen. Eine solche Vermutung ist psychologisch sinnvoll, da bei zunehmender Information über die Patienten eine anwachsende Ähnlichkeit der Urteile der Ärzte zu erwarten ist.

112. Skalen

Beim Skalen-Modus wurden pro Urteil 4 Faktoren extrahiert und zur Interpretation wie üblich varimaxrotiert. Darüber hinaus wurden die drei ermittelten Faktorenmatrizen auf ihre Ähnlichkeit hin überprüft.

Während die Anzahl signifikanter Faktoren sowie die Varianzstärke der Faktoren Hinweise auf den Grad der Differenziertheit der Urteile pro Modus liefern, geben diese Ähnlichkeitskoeffizienten Aufschluß über das Ausmaß der inhaltlichen Übereinstimmung der verglichenen Faktorenmatrizen. Diese Ähnlichkeitsüberprüfung erfolgt nach einer Methode von FISCHER und ROPPERT (1964), nach der beim Vergleich zweier Matrizen eine Transformationsmatrix ermittelt wird, die eine maximale Annäherung der einen Matrix an die andere, die invariante Matrix, herbeiführt. Über die theoretische Verteilung der nach dieser Methode errechenbaren Ähnlichkeitskoeffizienten ist nichts bekannt, so daß Angaben über die Signifikanz der Koeffizienten nicht möglich sind. GEBHARDT (1967) meint jedoch nach verschiedenen Ähnlichkeitsvergleichen, daß Koeffizienten bis etwa 0.4 oder 0.5 auf keinerlei Verwandtschaft der Faktorenstrukturen hinweisen, Koeffizienten unter 0.8 nur mit Vorsicht auf Faktorenstrukturähnlichkeit interpretiert werden dürfen, während Ähnlichkeitskoeffizienten ab 0.9 als deutlicher Hinweis auf eine Faktorenstrukturähnlichkeit und Koeffizienten in der Nähe von 1.0 als Hinweis auf annähernde Faktorenstrukturübereinstimmung gewertet werden können. Es ist jedoch zu beachten, daß die Höhe der Koeffizienten von der Anzahl der Variablen und von der Stichprobengröße abhängig ist. Die Koeffizienten wachsen mit Abfall der Variablenanzahl und Anstieg der Stichprobengröße.

Der Ähnlichkeitsvergleich der 3 Skalenfaktorenmatrizen erbrachte folgendes Ergebnis (Tabelle 1):

Urteil	2	3
1	.947	.949
2		.951

Tabelle 1 : Ähnlichkeitskoeffizienten aus dem Vergleich der 3 Skalenfaktorenmatrizen

Die Skalenfaktorenstruktur ist gegenüber Informationsart und -ausmaß fast invariant. Damit wird die Hypothese bestätigt, daß die durchschnittliche implizite Persönlichkeitstheorie, bzw. der kognitive Beurteilungsraum als deren Ausdruck die Skalenfaktorenstruktur gilt, vom Kenntnisumfang über die beurteilten Patienten beinahe unabhängig und somit sehr stabil ist. Diese Formulierung bedarf jedoch insofern einer Einschränkung, als anhand der prozentualen Faktorenvarianzanteile geringfügige Veränderungen zwischen den Urteilen nachgewiesen wurden. Die aufgewiesene Stabilität der durchschnittlichen impliziten Persönlichkeitstheorie betrifft daher hauptsächlich den Inhalt der Faktoren und weniger die Stärke der Faktoren oder die Stärke des Zusammenhangs der Skalen innerhalb eines Faktors.

Da die Skalenfaktorenmatrizen fast identisch sind, sollen hier lediglich die Faktoren aus dem 2. Urteil dargestellt und interpretiert werden. Die Faktoren sind auf den Betrag eins "normiert". Daher sind die Ladungen relativ gering. Bei der Interpretation der Faktoren werden daher Skalenladungen bis zur Höhe von .250 berücksichtigt. Skalen mit negativem Vorzeichen werden in der Darstellung invertiert, um die Interpretation zu erleichtern; das negative Vorzeichen wird jedoch beibehalten (Tabelle 2).

Faktor I	3)	-.458	geschlossen	-	offen
	9)	-.444	mißtrauisch	-	arglos
	1)	.379	mürrisch-gereizt	-	freundlich-ausgeglichen
	17)	-.339	resonanzunfähig	-	emotional gut ansprechbar
	13)	.336	gefühlarm	-	warmherzig
Faktor II	15)	.462	aggressiv	-	ängstlich zurückhaltend
	5)	-.434	antriebsgesteigert	-	antriebsgehemmt
	14)	.423	selbstbewußt	-	selbstunsicher
	10)	-.358	führend	-	sich unterordnend
	8)	-.327	heiter	-	traurig

Faktor III

12)	.501	sprunghaft	-	geordnet
2)	.490	weitschweifig	-	zielgerichtet
6)	-.376	ablenkbar	-	beharrlich
20)	.324	störbar	-	unbeeindruckbar
18)	.274	ratlos	-	entschieden
7)	-.265	gespannt	-	gelöst

Faktor IV

4)	-.579	differenziert	-	undifferenziert
11)	.567	sensibel	-	stumpf
16)	-.488	interessant	-	langweilig

Tabelle 2 : 4 rotierte Skalenfaktoren aus Urteil 2

Alle 4 Faktoren sind psychologisch plausibel.

Faktor I ist als "emotional verschlossen vs. emotional offen" interpretierbar. Im Pol "emotional verschlossen" werden offensichtlich auch paranoide Komponenten angesprochen.

Faktor II besteht aus den Polen "selbstsicher außenweltgerichtet vs. selbstunsicher zurückhaltend". Die Seite "selbstsicher außenweltgerichtet" enthält manische Aspekte.

Faktor III kann als "zerfahren vs. geordnet" interpretiert werden. Die Adjektive "ablenkbar vs. beharrlich" und "störbar vs. unbeeindruckbar" erweitern den Faktor in Richtung auf "emotional störbar vs. emotional ausgeglichen". Damit kommen diesem Faktor sowohl aktive als auch reaktive Komponenten zu. Im Pol "zerfahren" werden Schizophrenie-Aspekte angeschnitten.

Faktor IV enthält mit dem Adjektivpaar "differenziert vs. undifferenziert" kognitive und mit dem Adjektivpaar "sensibel vs. stumpf" emotionale Aspekte der Differenziertheit. Das Adjektivpaar "interessant vs. langweilig" in diesem Faktor hat insofern eine besondere Bedeutung, als es diesen Faktor klar als einen Bewertungsfaktor kennzeichnet:

Differenzierte und sensible Patienten sind für den explorierenden Arzt interessant, undifferenzierte und stumpfe langweilig.

Aus normalpsychologischer Sicht enthalten auch insofern die anderen Faktoren Bewertungsaspekte als der jeweilige, anormales Verhalten charakterisierende Pol negative und der normales Verhalten kennzeichnende Pol positive Bewertung impliziert. Eine solche Interpretation ist jedoch nicht zulässig, da das Polaritätenprofil speziell zur Beurteilung psychiatrisch Kranker entwickelt und auch dazu benutzt wurde. Eine bewertende Stellungnahme spielt demnach im durchschnittlichen Urteil der Psychiater nur dann eine Rolle, wenn sie Persönlichkeitseigenschaften einschätzen, die weniger eindeutig Komponenten einer psychischen Störung betreffen, sondern sich eher auf die Explorationssituation als soziale Situation und auf die Ergiebigkeit der kommunizierten Information auswirken. Die Exploration eines undifferenzierten und stumpfen Patienten ist eben langweilig. Dagegen erscheint die Exploration eines psychisch gestörten Patienten, der verschlossen, mißtrauisch, mürrisch-gereizt, resonanzunfähig und gefühlsarm ist (Faktor I) und der aufgrund dieser Eigenschaften die Explorationssituation ebenfalls schwierig gestaltet und wenig Information vermittelt, als eine psychiatrische Aufgabe, und entsprechend bleibt der Patient unbewertet.

Daß die bewertete Stellungnahme der Psychiater vorrangig im Zusammenhang mit der Explorationssituation und der Ergiebigkeit der vermittelten Information steht und sich weniger auf den beurteilten Patienten selbst bezieht, wird durch folgenden Sachverhalt bekräftigt:

Die einzigen der 21 Skalen, die in keinem der Faktoren eine Ladungshöhe von .250 erreichten, sind die beiden direkt personbezogenen, bewertenden Skalen "sympathisch vs. unsympathisch" und "unglaublich vs. ehrlich".

Dieses Ergebnis steht im Gegensatz zu vielen anderen (vgl. COHEN, 1969), anhand denen nachgewiesen wurde, daß sich bewertende Stellungnahmen insbesondere bei personbezogenen Eigenschaften auswirken. Diese Diskrepanz kann folgendermaßen erläutert werden: Die Urteiler in den verschiedenen Untersuchungen waren nur wenig, meist gar nicht psychiatrisch geschult. Die berufliche Aufgabe des Psychiaters ist es aber, Personen, insbesondere psychisch Kranke zu beurteilen und möglichst unabhängig von einer positiven oder negativen Haltung zu den Patienten Therapien einzuleiten.

Die Psychiater in dieser Untersuchung benutzten die beiden Skalen "sympathisch vs. unsympathisch" und "unglaublich vs. ehrlich" pro Urteil mit nur sehr geringer Var-

anz, d.h. sie differenzierten in ihrem bewertenden Urteil kaum zwischen den Patienten. Damit spielen bei ihnen im Gegensatz zu den Personen aus den verschiedenen Untersuchungen personbezogene, wertende Urteils-Komponenten tatsächlich nur eine geringe Rolle. Dieser Sachverhalt wird jedoch durch die 3 Modi-FA verstärkt, da diese Analyse-Methode Skalenvarianzunterschiede in voller Stärke berücksichtigt, während Korrelationen, auf denen eine übliche FA basiert, Varianzunterschiede fast ausgleichen. Dennoch kann die Aussage, im durchschnittlichen Urteil von Psychiatern spiele eine personbezogene, wertende Stellungnahme nur eine untergeordnete Rolle, beibehalten werden. Denn auch in der individuellen Urteilsstruktur tritt nur bei 2 der 5 Psychiater die Skala "sympathisch vs. unsympathisch" und nur bei einem die Skala "unglaublich vs. ehrlich" mit einer Ladung größer als .250 auf, und die Differenz zwischen den drei Urteilen ist bei keinem der 5 Ärzte weder im Sympathieausmaß noch im beurteilten Glaubwürdigkeitsgrad signifikant.

(Friedman Test; Sympathie: $\chi^2_1 = 2,50$, $p_1 > 0,05$; $\chi^2_2 = 1,03$, $p_2 > 0,05$;

$\chi^2_3 = 1,90$, $p_3 > 0,05$; $\chi^2_4 = 0,10$, $p_4 > 0,05$; $\chi^2_5 = 3,10$, $p_5 > 0,05$;

beurteilte Glaubwürdigkeit: $\chi^2_1 = 0,83$, $p_1 > 0,05$; $\chi^2_2 = 0,30$, $p_2 > 0,05$;

$\chi^2_3 = 0,40$, $p_3 > 0,05$; $\chi^2_4 = 0,10$, $p_4 > 0,05$; $\chi^2_5 = 3,33$, $p_5 > 0,05$

bei jeweils $df = 2$).

Aufgrund von Ergebnissen aus anderen Untersuchungen (z.B. SHRAUGER, 1967) ist bei zunehmender Information eine Erhöhung der positiven Bewertung zu erwarten.

Ob das geringe Ausmaß einer personbezogenen bewertenden Stellungnahme im durchschnittlichen Urteil der Psychiater möglicherweise auf rationale Kontrolle zurückzuführen ist, läßt sich natürlich nicht überprüfen.

113. Patienten

Es wurden pro Urteil aus Vergleichsgründen 4 Faktoren extrahiert, obwohl im ersten Urteil nur 3 Faktoren als bedeutsam gelten können. Wie beim Skalen-Modus wurden die 4 Faktoren zur Interpretation jeweils varimaxrotiert und die 3 Faktorenmatrizen auf ihre Ähnlichkeit hin überprüft (Tabelle 3).

Urteil	2	3
1	.779	.750
2		.907

Tabelle 3: Ähnlichkeitskoeffizienten aus dem Vergleich der
3 Patientenfaktorenmatrizen

Die Patientenfaktorenstruktur ist von Informationsausmaß und -inhalt nicht unabhängig. Ab 2. Urteil erreicht sie jedoch ein relativ hohes Maß an Stabilität. Damit ist die Hypothese bestätigt, daß sich das Urteil über die Patienten relativ schnell verfestigt, d.h. die innerhalb der inhaltlich varianten Beurteilungsstruktur lokalisierten Patiententypen verändern sich von 2. zu 3. Urteil inhaltlich nur noch wenig. Ob bei weiteren Urteilen noch entscheidende Verschiebungen auftreten, kann mit diesen Daten nicht überprüft werden, ist jedoch verhältnismäßig unwahrscheinlich (vgl. MEEHL, 1960).

Folgender Sachverhalt muß jedoch noch bei der Interpretation der Ähnlichkeitskoeffizienten berücksichtigt werden: Im ersten Urteil können nur 3 Faktoren als signifikant angesehen werden, was sicherlich z.T. die geringe Ähnlichkeit der Faktorenstrukturen aus erstem und den beiden folgenden Urteilen mitbedingt. Eine vorsichtige Interpretation in Bezug auf Veränderungen zwischen 1. und 2. Urteil lautet daher folgendermaßen: Die im Urteil der Psychiater implizierten Patiententypen ändern sich vom 1. zum 2. Urteil. Diese Veränderung geht zu Lasten von a) einer Vergrößerung der Anzahl von Patiententypen, womit die Ausdifferenzierungshypothese bestätigt ist, und b) zu Lasten einer inhaltlichen Veränderung der Patiententypen. Die Aussage, die Patiententypen verändern sich von 2. zu 3. Urteil inhaltlich nur noch wenig, kann beibehalten werden, da in beiden Urteilen 4 Faktoren als signifikant zu betrachten sind. Die Ähnlichkeit zwischen 2. und 3. Urteil bezieht sich auf den Inhalt zweier Faktorenmatrizen mit gleicher Anzahl signifikanter Faktoren.

Von Interesse ist, ob sich die Patientengruppen N, S und Z bezüglich des Ähnlichkeitsausmaßes zwischen 1. und 2., 2. und 3. und 1. und 3. Urteil unterscheiden. Eindeutige Hypothesen bestehen bei dieser Fragestellung nicht. Es traten keine signifikanten Differenzen auf, d.h. die laut Außenkriterien festgelegten Patientengruppen bewirken keine differentiellen Unterschiede bezüglich Stabilität der Beurteilung (Kruskal-Wallis Test: $H_{1/2} = 1,94$, $p > 0,05$; $H_{2/3} = 0,02$, $p > 0,05$; $H_{1/3} = 1,28$, $p > 0,05$;

bei jeweils $df = 2$). Als Meßwertträger gingen die jeweiligen Ähnlichkeitskoeffizienten pro Patient in die Berechnungen ein.

Zur Interpretation der Patientenfaktoren pro Urteil können die von den jeweils behandelnden Ärzten getroffenen Diagnosen als Außenkriterium herangezogen werden. Diese Diagnosen stehen in der folgenden Zusammenstellung hinter den Ladungen. Da die Faktoren gleichgewichtig sind, wurden sich entsprechende Faktoren nebeneinandergestellt. Sich entsprechende Faktoren können umgepolt sein. Es sind nur Ladungen größer, gleich .25 aufgeführt (Tabelle 4).

Urteil	1	2	3
Faktor			
I	7) -.95 S	7) .90 S 10) .32 S	7) .75 S
II	1) .62 N 15) .44 Z 2) .38 N 11) .27 Z	1) -.56 N 11) -.53 Z 15) -.36 Z 9) .34 S 10) .27 S	1) -.70 Z 9) .41 S 1) -.34 N 10) .29 S
III	10) .44 S 11) .43 Z 1) -.43 N 6) .42 S 12) .34 Z	10) -.51 S 12) -.43 Z 6) -.39 S 11) -.30 Z 4) -.29 N 8) .27 S	1) .59 N 12) -.51 Z 8) .31 S 4) -.30 N 6) -.28 S
IV	14) .49 Z 3) .48 N 9) .39 S 12) -.38 Z 5) .31 N 4) .30 N	2) .58 N 1) .34 N 11) -.29 Z 3) .28 N 9) .27 S 13) .26 Z 10) .25 S	13) .39 Z 1) .35 N 5) .35 N 10) .33 S 3) .32 N 14) .31 Z 4) .29 N 7) -.28 S 2) .26 N

Tabelle 4: 4 rotierte Patientenfaktoren pro Urteil

Die Interpretation der Faktoren anhand des Außenkriteriums "Diagnose" gestaltet sich relativ schwierig. Lediglich der jeweils erste Faktor ist eindeutig, jedoch handelt es sich jeweils um einen recht spezifischen Faktor. Das wird insbesondere im ersten Urteil deutlich, in den beiden folgenden Urteilen verliert dieser Faktor an Spezifität. Dieser Verlust ist als Hinweis auf ein Herauskristallisieren eines entsprechenden Patiententyps zu werten. Patient 7 war ein akut paranoider Schizophrener, Patient 10 befand sich am Anfang eines entsprechenden Prozesses.

Die weiteren Faktoren pro Urteil beinhalten jeweils Patienten mit mindestens 2 unterschiedlichen Diagnosen laut Außenkriterium. Faktor II des ersten Urteils ist ein unipolarer Faktor, auf den nur Z und N hoch laden. Dieser Faktor erweitert sich im 2. und 3. Urteil zu jeweils einem bipolaren Faktor, in dem Patienten der Diagnosegruppe S den Gegenpol bilden. Diese Polarisierung, N- und Z-Diagnosen vs. S-Diagnosen, macht wiederum deutlich, daß ab 2. Urteil eine Verschärfung der Faktorenklarheit bzw. Faktorenbestimmtheit eintritt. Die Polarität N- und Z-Diagnosen vs. S-Diagnosen ist psychopathologisch sinnvoll, wenn man berücksichtigt, welche Beurteilungsskalen zur Verfügung standen. Andererseits ist die Bipolarität dieser beiden Faktoren sowie die des Faktors III im 3. Urteil, die ebenfalls sehr eindeutig ist, insofern besonders interessant, als sie im Rahmen der Patientenanalyse belegt, daß Patiententypen im psychologischen Urteil der Ärzte nicht nur Charakteristika kategorialer Klassen haben (wie z.B. Patiententyp bzw. Faktor I), sondern auch die relationaler Klassen. Die Patiententypen werden durch die Gegenpole einer Patientendimension bestimmt; ein Typ ist gewissermaßen Anti-Typ des anderen.

Mit Ausnahme des Faktors III im ersten Urteil, der aus den Polen N- vs. S- und Z-Diagnosen besteht, bilden die weiteren Faktoren in allen 3 Urteilen ein unklares Gemisch aus Patienten der 3 Diagnosegruppen. Eine eindeutige Interpretation allein aufgrund des Außenkriteriums "Diagnose" ist nicht möglich. Der Versuch, die Interpretation der Faktoren nicht anhand der als Außenkriterium geltenden Diagnosen vorzunehmen, sondern anhand der von den Ärzten pro Patient jeweils am häufigsten getroffenen Diagnosen, erleichtert nicht die Interpretation. Er soll daher hier unterlassen werden. Die Schwierigkeit der Interpretation der Faktoren anhand der Diagnose ist ein übliches Problem (vgl. ZUBIN, 1967). Verantwortlich für die Schwierigkeiten können a) aus psychiatrischer Sicht ungeeignete Beurteilungsskalen oder b) bei geeigneten Beurteilungsskalen das damit

auch im Urteil der Psychiater ungeeignete Diagnosegruppensystem sein. Zu der letzten der beiden Schlußfolgerungen neigen die meisten Autoren.

Die aufgrund der Ladungshöhen errechneten Kommunalitäten repräsentieren im Hauptkomponentenmodell den Teil der Varianz einer Variable, der mit den die Gesamtvarianz aufklärenden Faktoren geteilt wird. Die Einheitsvarianz einer jeden Variable wird als gemeinsame Varianz betrachtet. Der Begriff "Kommunalität" ist beim Hauptkomponentenmodell dann insofern irreführend, als es zu Faktoren kommen kann, in denen nur eine einzige Variable hoch lädt. Solche Faktoren als spezifische Faktoren zu bezeichnen, wäre dann wiederum nicht korrekt, da ja die Einheitsvarianz einer jeden Variable als gemeinsame Varianz gilt (zur Problematik dieses Modells, das nur zu deskriptiven Faktoren führt, siehe u.a. ÜBERLA, 1968). Jedenfalls kann eine Variable eine hohe Kommunalität erhalten, wenn sie alleine einen Faktor aufspannt oder komplex ist, d.h. in mehreren Faktoren relativ hoch geladen auftritt. Dagegen weist eine geringe Kommunalität einer Variable darauf hin, daß die Varianz dieser Variable nur zu einem geringen Teil durch die extrahierten, die Gesamtvarianz erfassenden Faktoren aufgeklärt wird. In der 3-Modi-FA ergeben sich noch dadurch weitere Schwierigkeiten, daß alle Faktoren auf den Betrag Eins "normiert" sind, womit die unterschiedliche Varianzstärke der Faktoren ausgeschaltet wird. Das hat zwar den Vorteil, daß alle Faktoren in der Kermatrix gleichgewichtig und somit untereinander vergleichbar sind, wirkt sich jedoch insofern nachteilig aus, als bei Faktoren, die varianzstark und damit mit vielen Variablen besetzt sind, die Ladungshöhe dieser Variablen durch die Stauchung der Faktoren vermindert wird. Entsprechend sind die Kommunalitäten dieser Variablen relativ gering. Andererseits werden Faktoren mit nur wenigen Variablen gestreckt und die Höhe der Ladungen der Variablen sowie deren Kommunalitäten vergrößert. Und Variablen, die einen Faktor alleine ausmachen, erhalten eine hohe Ladung und demgemäß eine hohe Kommunalität. Dieses Vorgehen berührt zwar nicht den Ähnlichkeitsvergleich von Faktorenstrukturen, die Interpretation der in der 3-Modi-FA errechneten Kommunalitäten ist jedoch sehr schwierig. Die Interpretation der Kommunalität einer Variablen dürfte an sich nur unter Berücksichtigung der Art des bzw. der Faktoren, in dem bzw. denen die Variable eindeutig lädt, vorgenommen werden. Faßt man jedoch die bisherigen Überlegungen zusammen, dann kann eine hohe Kommunalität einer Variablen noch am ehesten als Ausdruck für eine "Urteils-Bestimmtheit" dieser Variable im Verhältnis zu Variablen mit niedriger Kommunalität angesehen werden, da sie entweder alleine einen Faktor aufmacht oder komplex ist. Die Patientenkommu-

litäten gelten somit im folgenden als Index für das Ausmaß der Bestimmtheit der Patienten.

Die Höhe von Ranginterkorrelationen zwischen den Kommunalitäten aus 1., 2. und 3. Urteil geben an, in welchem Ausmaß die Information zwischen den Urteilen Einfluß auf den relativen Bestimmtheitsgrad der Patienten nimmt. Eine hohe positive Korrelation zeigt an, daß Patienten, die in einem Urteil einen hohen bzw. niedrigen Bestimmtheitsgrad erreichten, einen solchen auch in anderen Urteilen erhielten (Tabelle 5).

Urteil	2	3
1	.42	.42
2		.72

Tabelle 5: Beziehung zwischen den Patienten-Kommunalitäten aus den 3 Urteilen (Rangkorrelationskoeffizienten) $r_s \geq .43$, $p \leq 0,05$; $r_s \geq .63$, $p \leq 0,01$

Auch anhand dieser Daten ist eine Stabilisierung des Urteils ab 2. Urteil festzustellen. Dieses Ergebnis wird natürlich vom Ergebnis zum Faktorenstrukturvergleich impliziert. Die folgende Tabelle 6 enthält für Urteil 1, 2 und 3 die pro Diagnosegruppe (laut Außenkriterium) summierten Kommunalitäten der Patienten.

	Diagnose	N	S	Z
Urteil 1		1,26	1,65	1,09
2		1,20	1,80	0,99
3		1,16	1,61	1,22
Σ		3,63	5,06	3,30

Tabelle 6: Summe der Patientenkommunalitäten pro Diagnosegruppe und Urteil

Die Änderung des Ausmaßes der Urteilsbestimmtheit von 1. bis 3. Urteil pro Diagnosegruppe ist in keinem der 3 Fälle signifikant (Friedman Test: $\chi^2_N = 1,2$; $\chi^2_S = 0,00$; $\chi^2_Z = 1,6$ bei jeweils $df = 2$; p jeweils $> 0,05$). Die Daten weisen bezüglich des Ausmaßes der Urteilsbestimmtheit auf Unterschiede zwischen den Diagnosegruppen hin. Schizophrene scheinen pro Urteil bestimmter eingeschätzt worden zu sein. Die Unterschiede sind jedoch nicht signifikant (Kruskal-Wallis Test; $H_1 = 0,05$; $H_2 = 1,04$;

$H_3 = 0,66$ bei jeweils $df = 2$, p jeweils $> 0,05$).

Die Höhe der - ohne Berücksichtigung der Vorzeichen - gemittelten Korrelationen zwischen der durch einen Arzt getroffenen Beurteilung eines realen Patienten und den durch denselben Arzt aus der Vorstellung heraus vorgenommenen Beurteilungen mehrerer typischer Patienten, die sich in ihrer Symptomatik stark voneinander unterscheiden, kann als Maß der beurteilten stereotypenhaften Symptom- bzw. Syndrommannigfaltigkeit interpretiert werden. Eine hohe durchschnittliche Korrelation zeigt an, daß sich der Arzt bei der Beurteilung des Patienten an einer Vielzahl von Symptomen als Beschreibungstereotypen orientiert hat. Die Einschätzung des Patienten erfolgte somit im Sinne der Stereotype differenziert, die individuelle psychologische Persönlichkeitsstruktur des Patienten wäre jedoch wenig berücksichtigt. Eine hohe Korrelation ist damit gewissermaßen auch Indiz dafür, daß die Symptome eines Patienten im Sinne der Stereotype vom Arzt erkannt und psychologisch kategorisierend beschrieben wurden. Entsprechend verliert der Patient seine psychologische oder psychopathologische Individualität.

Zur Bestimmung der stereotyphaften Symptomvielfalt werden pro Arzt zwischen den jeweiligen 3-fachen Beurteilungen eines jeden Patienten und den Beurteilungen der 9 typischen aus der Vorstellung heraus eingeschätzten Patienten Q-Korrelationen errechnet. Die dieserart ermittelten Daten geben Aufschluß über das Ausmaß der stereotypen Beurteilung eines Arztes pro Patient und Urteil. Bei der Berechnung intraindividueller Unterschiede pro Arzt und interindividueller Differenzen zwischen den Ärzten bezüglich der Höhe der Korrelationen fungieren jeweils Patienten als Meßwertträger. Die Aussagen aufgrund solcher Berechnungen betreffen die Ärzte. Verwendet man dagegen die Ärzte als Meßwertträger, dann sind Aussagen über den beurteilten stereotypen Symptomreichtum der Patienten möglich, und es können entsprechend intra- und interindividuelle Unterschiede bezüglich der Patienten berechnet werden. Nicht-parametrische Berechnungen, in die gleichzeitig beide Meßwertträgerarten eingehen, erlauben keine differentiellen Aussagen in Bezug auf Ärzte oder Patienten und sind wegen einer möglichen Interaktion zwischen den beiden Meßwertträgern nicht sinnvoll.

Bevor auf die die Patienten betreffenden differentiellen Aspekte eingegangen wird, die innerhalb dieses Abschnittes ihren Platz haben, soll noch folgendes allgemeines Ergebnis referiert werden.

Der nach z-Transformation errechnete Mittelwert aller Korrelationen beträgt $r = .30$. Diese mittlere Korrelation ist nicht hoch. Die Einschätzung der Patienten orientierte

sich somit nur mäßig an stereotyper Symptomatik.

Bezüglich der die Patienten betreffenden differentiellen Aspekte ist es von Interesse, ob sich a) Unterschiede zwischen den 3 Urteilen und b) pro Urteil Unterschiede zwischen den Diagnosegruppen im Ausmaß des beurteilten stereotypen Symptomreichtums ergeben. Dabei besteht bezüglich a) die Hypothese, daß mit zunehmender Information über die Patienten deren Symptomatik von den Ärzten klarer erfaßt und im Sinne der Stereotypen beurteilt wird. Entsprechend ist eine Erhöhung der durchschnittlichen Korrelationen von Urteil 1 nach Urteil 3 zu erwarten. Bezüglich b) besteht keine gezielte Hypothese.

Die pro Urteil und Diagnosegruppe über die Ärzte gemittelten Korrelationen (Mittlung nach z-Transformation) sind in Tabelle 7 aufgeführt.

Diagnosegruppe	N	S	Z
Urteil 1	.27	.35	.27
2	.29	.34	.30
3	.28	.32	.29

Tabelle 7: Mittel der Korrelationskoeffizienten pro Diagnosegruppe und Urteil

Bei der Überprüfung der Hypothesen a) und b) sollten Ärzte als Meßwertträger fungieren, was jedoch wegen des geringen N von nur 5 statistisch nicht sinnvoll erscheint. Daher wird folgender Weg beschritten: Die über die Ärzte pro Patient berechnete durchschnittliche Korrelation gilt als Maß für die stereotype Symptomvielfalt des Patienten. Nach entsprechenden Berechnungen pro Patient und Urteil kann mit Hilfe des Friedman Test Hypothese a) überprüft werden. Die Unterschiede sind nicht signifikant ($\chi^2 = 1,29$ bei $df = 2$, $p > 0,05$) und damit Hypothese a) widerlegt. Art und Ausmaß der über die Patienten gewonnenen Information beeinflußt somit nicht signifikant ihren beurteilten stereotypen Symptomreichtum. Informationsart und -ausmaß haben auch keinen entscheidenden Einfluß auf das Verhältnis der Patienten zueinander bezüglich dieses Stereotypmaßes. Die Rangkorrelationen zwischen dem Maß aus 1. und 2. Urteil, sowie zwischen 2. und 3. Urteil sind jeweils hoch signifikant; sie betragen $r_{1/2} = .743^{**}$ und $r_{2/3} = .900^{**}$. Die Korrelationen machen jedoch deutlich, daß die Konstanz des Verhältnisses der Patienten zueinander von der Information aus Situation IV und V weniger beeinträchtigt wird als von der Information aus Situation II und III. Gleichzeitig wird auch anhand

dieser Daten eine Stabilisierung des Urteils ab 2. Urteil nachgewiesen.

Bei Frage b), nämlich Unterschiede zwischen den Diagnosegruppen pro Urteil, traten ebenfalls keine signifikanten Differenzen auf. (Kruskal-Wallis Test: $H_1 = 4,16$; $H_2 = 3,02$; $H_3 = 3,38$ bei jeweils $df = 2$; p jeweils $> 0,05$). Die Diagnosegruppen unterscheiden sich somit nicht bezüglich ihres beurteilten Symptomreichtums.

Zwischen beurteiltem stereotypen Symptomreichtum der Patienten und ihrer Urteils-Bestimmtheit (Kommunalitätshöhe) besteht kein Zusammenhang. Die Rangkorrelation pro Urteil ist jeweils nicht signifikant ($r_1 = -.254$; $r_2 = .147$; $r_3 = .394$).

Zusammenfassend sind Urteilsbestimmtheit sowie Urteilsstereotypie nur minimal von den Bedingungen Informationsausmaß und -inhalt sowie Diagnosegruppe abhängig, d.h. das psychologische Urteil erweist sich gegenüber diesen Bedingungen als relativ invariant. Dieses Ergebnis kann im Zusammenhang mit den eingangs aufgeführten Überlegungen zum psychologischen Urteilsprozeß gesehen werden. Das psychologische Urteil verläuft offensichtlich relativ wenig beeinflussbar auf einer niederen Bewußtseinsbene.

114. Beurteiler

Pro Urteil wurden 2 Faktoren extrahiert, eindeutig bedeutsam ist jedoch jeweils nur ein Faktor. Mit diesem Vorgehen wird bezweckt, mögliche Differenzen zwischen den Beurteilern besser in den Griff zu bekommen. GITIN (1970) hält es für sinnvoll, mehrere Beurteilerfaktoren zu extrahieren. Die zwei Faktoren wurden jeweils varimaxrotiert und die drei Faktorenmatrizen auf Ähnlichkeit hin überprüft (Tabelle 8).

Urteil	2	3
1	.752	.573
2		.888

Tabelle 8: Ähnlichkeitskoeffizienten aus dem Vergleich der 3 Beurteilerfaktorenmatrizen

Die Ähnlichkeitskoeffizienten müssen als etwas unterschätzt angesehen werden, da der in die Vergleiche eingegangene Faktor nicht genügend abgesichert ist.

Auch anhand dieser Ähnlichkeitsvergleiche wird die Stabilisierung des Urteils ab 2. Urteil nachgewiesen. Die Information aus Situation IV und V trägt nur noch geringfügig zu einer Veränderung der Beurteilertypen bei.

Der Vergleich mit den entsprechenden Ergebnissen unter Patienten-Modus zeigt, daß vom 1. zum 3. Urteil eine relativ starke Änderung der Beurteilertypen erfolgt (Ähnlich-

keitskoeffizient .573). Bei den Patiententypen beträgt der entsprechende Koeffizient .750. Die Beurteilertypen scheinen bei viel Information stärker änderungsanfällig zu sein als die Beurteilten. Es muß hier jedoch einschränkend berücksichtigt werden, daß der Effekt des jeweils 2., nicht eindeutig signifikanten Faktors, auf diese Ähnlichkeitsvergleiche nicht ausreichend abgeschätzt werden kann. Das genannte Ergebnis ist möglicherweise ein Artefakt.

Auf die Darstellung der Arztfaktoren soll verzichtet werden. Es ist lediglich hervorzuheben, daß der zweite Faktor jeweils durch Arzt 1 oder 4 bestimmt wird, während sich die restlichen Ärzte relativ gleichmäßig auf den jeweils ersten Faktor verteilen.

115. Kernmatrix

Die Werte in der Kernmatrix geben Stärke (Wertgröße) und Richtung (Vorzeichen) der Beziehung zwischen Skalen- und Patientenfaktoren pro Arztfaktor an. Da pro Urteil in allen Modi-Matrizen bipolare Faktoren auftreten, ist die Richtung des Zusammenhangs zwischen Skalen- und Patientenfaktoren nicht eindeutig bestimmt. Daher werden die Werte in der Kernmatrix lediglich als Gewicht der Skalenfaktoren bei der Bestimmung der Patientenfaktoren interpretiert; die Vorzeichen bleiben unberücksichtigt. Die umgekehrte Interpretation, d.h. die Werte in der Kernmatrix geben das Gewicht der Patiententypen bei der Bestimmung der Skalenfaktoren an, ist natürlich auch möglich, da die Werte in der Kernmatrix durch Multiplikation der Skalen-, Patienten- und Arztfaktoren zustandekommen. Die erste der beiden Interpretationen soll jedoch weiterverfolgt werden, da sie psychologisch plausibler ist und besser in die Thematik der Untersuchung paßt.

Ein metrischer Vergleich der Kernmatrizen aus den 3 Urteilen ist nicht möglich, da laut Faktorenstrukturvergleich die jeweiligen Modusmatrizen nicht identisch und die Kernmatrizen somit inhaltlich unterschieden sind. Auch ein interpretativer Vergleich soll unterlassen werden. Stattdessen wird hier exemplarisch die Kernmatrix des 2. Urteils dargestellt (Tabelle 9).

Arzt- faktor	Patienten- faktor	Skalen- faktor	1	2	3	4
1	1		-6,91	7,30	12,62	-3,37
	2		-10,66	-5,81	0,30	-6,75
	3		4,22	-1,90	1,04	4,16
	4		-2,08	-11,56	-1,41	2,55
2	1		-6,66	4,19	3,46	-4,39
	2		-6,03	-5,44	3,86	-2,92
	3		-0,67	-6,72	7,16	2,51
	4		5,98	0,19	-8,12	-0,96

Tabelle 9 : Kernmatrix des 2. Urteils

Aus den dargestellten Werten wird deutlich, daß die Skalenfaktoren zur Beurteilung der verschiedenen Patiententypen von recht unterschiedlicher Bedeutung sind. Das gilt für beide Arzttypen. Zwischen den Arzttypen bestehen bezüglich der Gewichtigkeit der Skalenfaktoren für die verschiedenen Patiententypen recht beträchtliche Unterschiede. Das wird bei Patiententyp 4 besonders deutlich. Für Arzttyp 1 ist Skalenfaktor 2 bei der Beurteilung dieses Patiententyps am wichtigsten, für Arzttyp 2 jedoch am unwichtigsten. Diese Unterschiede zwischen den Arzttypen sollten jedoch nicht überbewertet werden, da der 2. Arzttyp laut Eigenwerthöhe nicht eindeutig signifikant ist.

Eine allgemeine Interpretation dieser Kernmatrix ist schwierig, da nicht alle Patiententypen aufgrund des Außenkriteriums "Diagnose" ausreichend eindeutig bestimmt werden konnten. Die Kernmatrix kann nun aber dazu benutzt werden, über die Gewichtung der Skalenfaktoren für die Patiententypen die Patiententypen näher zu bestimmen. Nur beispielhaft sei auf den Patiententyp 1, es handelt sich um 2 Schizophrene, hingewiesen. Für Arzttyp 1 spielt bei der Beurteilung dieses Patiententyps die Beurteilungsdimension 3 "zefahren vs. geordnet" eine besonders große Rolle. Arzttyp 2 gewichtet dagegen Skalenfaktor 1 "emotional verschlossen vs. emotional offen" am stärksten, Skalenfaktor 3 ist wenig bedeutend. Auf eine weitere Interpretation dieser Kernmatrix soll verzichtet werden. Die beiden Kernmatrizen aus Urteil 1 und 3 befinden sich im Anhang III.

Für die Thematik dieser Untersuchung aufschlußreicher ist die Interpretation der Kernmatrix als kognitiver Beurteilungsraum, in dem die Patiententypen lokalisiert sind. Dabei ist jeder zu einem Arzttyp gehörende Teil einer Kernmatrix als ein solches kognitives System zu betrachten.

Eingangs war zwischen relationalen und kategorialen (konjunktiven und disjunktiven) Klassen unterschieden worden. Relationale Klassen werden durch alle Beurteilungsdimensionen bestimmt; sie unterscheiden sich bezüglich der Ausprägung auf den einzelnen Beurteilungsdimensionen. Kategoriale Klassen werden durch Merkmale bzw. Dimensionen definiert, die sich zwischen den Klassen unterscheiden, wobei die Ausprägung der Merkmale bzw. Dimensionen unberücksichtigt bleibt. Eine eindeutige Interpretation der Kernmatrizen unter dem Gesichtspunkt verschiedener formaler Klassen ist nicht möglich, da kategoriale Klassen dichotome Werte in den Kernmatrizen verlangen, d.h. eine kategoriale Klasse wird durch Merkmale bzw. Dimensionen entweder bestimmt oder nicht bestimmt. Ein entsprechender Versuch soll jedoch unternommen werden, wobei die Werte in den Kernmatrizen als Null ($-1,99$ bis $1,99$) bzw. als Nicht-Null (über $-2,0$ bzw. $2,0$) interpretiert werden. Lediglich Patiententyp 1 innerhalb Arzttyp 1 ist in diesem Sinne eine kategoriale Klasse, da allein dieser Patiententyp durch das Merkmal bzw. die Dimension 3 "z erfah ren vs. geordnet" bestimmt wird. Die drei weiteren Patiententypen dieses Arztfaktors sind relationale Klassen, sie werden alle durch die drei restlichen Dimensionen näher bestimmt. Bei Arzttyp 2 treten nur relationale Klassen auf.

Die Kernmatrizen können jedoch noch in anderer Hinsicht interpretiert werden: Die Anzahl der Skalenfaktoren, die für die Bestimmung eines Patiententyps relevant ist, d.h. die entsprechenden Werte in der Kernmatrix sind von Null verschieden, kann im Sinne der Differenziertheit der Beurteilung dieses Patiententyps interpretiert werden. Ein Patiententyp, der nur durch einen Skalenfaktor bestimmt wird, wird wenig differenziert beurteilt. Ein solcher Sachverhalt tritt z.B. in Urteil 1 bei beiden Arzttypen auf; Patiententyp 3 wird nur jeweils von einem Skalenfaktor charakterisiert. Das Ausmaß der Beurteilungsdifferenziertheit eines jeden Patiententyps ist natürlich abhängig von der Anzahl signifikanter Patiententypen. In Urteil 1 sind laut Eigenwertanalyse entsprechend auch nur 3 Patientenfaktoren eindeutig signifikant. Demgemäß kommt in Urteil 2 und 3, in dem jeweils 4 Patientenfaktoren als signifikant anzusehen sind, ein entsprechend eindeutiger Sachverhalt nicht vor.

Anhand der Anzahl von Nullwerten pro Kernmatrix aus Urteil 1, 2 und 3 soll nun die schon mit Hilfe der Eigenwerte nachgewiesene Ausdifferenzierung der Patientenfaktoren-

struktur nochmals aufgezeigt werden. Arzttypen werden dabei nicht berücksichtigt, sämtliche Null-Werte werden pro Kernmatrix summiert. Eine derartige Überprüfung der Ausdifferenzierung der Patientenfaktoren ist psychologisch aufschlußreicher als die entsprechende Überprüfung anhand der Eigenwerte. Denn die nachgewiesene Stabilität der Skalenfaktorenstruktur erlaubt die Interpretation, daß im inhaltlich konstanten Beurteilungsraum die sich ausdifferenzierenden Patiententypen mit einer zunehmenden Anzahl von Beurteilungsdimensionen belegt werden.

Gelten wiederum Werte zwischen -1,99 und 1,99 als Nullwerte, dann ergibt sich von Urteil 1 bis 3 folgende Verteilung der Null-Wert-Häufigkeiten: 12, 7 und 4. Noch eindrucksvoller ist die Verteilung, wenn Werte zwischen -0,99 und 0,99 als Null-Werte gesetzt werden: 9, 4 und 0.

12. Psychiatrisches Urteil

Bei der Analyse der psychiatrischen Diagnosen ergaben sich aus folgenden Gründen Schwierigkeiten:

- a) Die Ärzte gaben nicht nur exakte Diagnosen ab; häufig erfolgten Beschreibungen des klinischen Zustandsbildes oder es wurden gar keine Angaben gemacht.
- b) Z.T. wurden bei einem Urteil mehrere Diagnosen zu einem Patienten getroffen, manchmal entsprechend ihrer Wahrscheinlichkeit gewichtet.
- c) Die Nomenklatur der Ärzte unterschied sich.

Die Angaben mußten daher zunächst inhaltsanalytisch aufbereitet werden. Dabei wurden folgende Regeln beachtet:

- 1) Es wurden nur Grob-Diagnosen - insbesondere die Diagnosen N, S und Z - berücksichtigt, da bei Beachtung differenzierterer Diagnosegruppen vor allem wegen der unterschiedlichen Nomenklatur der Ärzte erhebliche Kategorisierungsschwierigkeiten aufkamen.
- 2) Unspezifische Beschreibungen wurden der Kategorie "Beschreibungen" (B) zugeordnet. Als unspezifische Beschreibungen gelten auch die Diagnosen "Himorganischer Abbau" und "Alkoholabusus". Eine solche Kategorisierung erscheint gerechtfertigt, da bei einigen Patienten diese Diagnosen als Zweit-Diagnosen tatsächlich zutrafen. Für diese Diagnosen eigene Kategorien bereitzustellen, hätte den Nachteil, daß diese Diagnosen dann unter Validitätsgesichtspunkten als Falsch-Diagnosen kategorisiert würden. Als "Beschreibung" wurde auch die unspezifische Diagnose "Psychose" gewertet. Geht aus den Be-

beschreibungen eindeutig hervor, welche Diagnosegruppe gemeint ist, wird die Beschreibung als Diagnose betrachtet und entsprechend kategorisiert. Z.B. gilt die Beschreibung "manisches Zustandsbild, Paranoide Psychose" als Diagnose "Schizophrenie".

Keine Angaben wurden als Beschreibung kategorisiert, da eine Beschreibung bei fehlender Diagnosekenntnis als Verlegenheitsangabe anzusehen ist.

3) Bei mehreren, bezüglich ihrer Zutreffens-Wahrscheinlichkeit vom Arzt ranggeordneten Diagnosen geht diejenige mit der größten Zutreffens-Wahrscheinlichkeit in die Kategorisierung ein.

Von mehreren, nicht ranggeordneten Diagnosen wurde die erste Diagnose berücksichtigt. Dieser Entscheidung liegt die Überlegung zugrunde, daß die vom Arzt erstgenannte Diagnose für ihn vermutlich auch die wahrscheinlichste war.

Die im folgenden dargestellten Ergebnisse zum psychiatrischen Urteil sind untergliedert in Ergebnisse zur Stabilität und Validität.

Als Meßwertträger finden Beachtung die beurteilten Patienten oder die beurteilenden Ärzte, nicht jedoch die Beurteilungswiederholungen wegen des zu geringen N von nur 3.

Die Ergebnisse zur Stabilität und Validität der Urteile sind jeweils differenziert nach allgemeinen Gesichtspunkten, bei denen 2 oder alle 3 Meßwertträger zusammengefaßt sind, und in differentielle Aspekte. Unter differentiellen Aspekten werden intra- und interindividuelle Unterschiede bei den Ärzten und bei den Patienten dargestellt, wobei im ersten Fall Patienten und im zweiten Fall Ärzte als Meßwertträger fungieren.

Als jeweils wichtigster Ergebnisteil gelten die Differenzen zwischen den drei Urteilen.

Unter "allgemeinen Gesichtspunkten" können die Daten wegen der Interaktion der Meßwertträger nicht einer statistischen Differenzüberprüfung unterzogen werden. Bei den differentiellen Aspekten sind - außer bei Zusammenfassung der Daten über alle drei Urteile - statistische Überprüfungen prinzipiell möglich, da in diesen Fällen nur eine Art von Meßwertträgern in die Berechnung eingeht. Wegen der häufig geringen Anzahl von Meßwertträgern sind jedoch einige der Überprüfungen statistisch nicht mehr sinnvoll und werden daher unterlassen.

Die Darstellung der Ergebnisse zur Stabilität und Validität unter allgemeinen Gesichtspunkten und intraindividuellen Differenzen orientiert sich an folgendem Schema:

- 1) Veränderung der Werte vom ersten bis dritten Urteil insgesamt
- 2) Die 1) entsprechenden Veränderungen pro Diagnosegruppe
- 3) Unterschiede zwischen den Diagnosegruppen insgesamt
- 4) Die 3) entsprechenden Unterschiede pro Urteil
- 5) Bei mehr als einem Meßwert interessiert - für alle drei Urteile insgesamt -
das Verhältnis der Meßwerte zueinander
- 6) Das 5) entsprechende Verhältnis pro Urteil

Es werden im folgenden nicht alle der hier genannten Ergebnisteile aufgeführt werden. Ob die Darstellung erfolgt, kann von statistischer Überprüfbarkeit der Daten und Relevanz in Bezug auf die Problemstellung abhängig sein.

121. Stabilität

Der Vergleich der Stabilität des psychiatrischen mit der des psychologischen Urteils gestattet die Überprüfung der Hypothese, daß sich das psychologische Urteil schneller stabilisiert als das psychiatrische.

Da im psychiatrischen Urteil nur Patiententypen zu erfassen sind, muß der Verlaufsvergleich die Patiententypen aus dem psychologischen Urteil betreffen. D.h., verglichen wird der Verlauf der psychologischen Sicht der Patienten (Patiententypen) mit dem der psychiatrischen Sicht der Patienten (Diagnosegruppen). Änderungen schlagen sich im psychologischen Urteil in unterschiedlichen Patiententypen, im psychiatrischen Urteil in Diagnosewechsel von Patienten nieder.

Die folgenden Tabellen 10, 11 und 12 beinhalten die Anzahlen der - von 1. zu 2., von 2. zu 3. und von 1. zu 3. psychiatrischen Urteil - übereinstimmenden und sich ändernden Diagnosen der Patienten. Es handelt sich um über die 5 Ärzte summierte Werte. Außer den Beschreibungen (B) und den N-, S- und Z-Diagnosen finden auch die eigens kategorisierten Diagnosen Psychopathie (Ps), endoreaktive Depression (enre D) und Entwicklungskrise (E) Berücksichtigung.

	B	N	S	Z	PS	ENRE D	E	
B	13 (47 %)	10	2	3	3			31
N	1 (92 %)	13						14
S	2		16 (90 %)	1			1	20
Z		1	1	5 (71 %)				7
PS	2							2
ENRE D								
E							1 (100 %)	1
	18	24	19	9	3		2	75

Tabelle 10: Anzahl der übereinstimmenden und sich ändernden Diagnosen zwischen Urteil 1 und 2. Nähere Erläuterung im Text.

	B	N	S	Z	PS	ENRE D	E	
B	2 (11 %)	3	1	8	1	3		18
N	1 (62 %)	15	5	2		1		24
S			16 (84 %)	3				19
Z		1	2	6 (67 %)				9
PS		1	1		1 (33 %)			3
ENRE D								
E			2					2
	3	20	27	19	2	4		75

Tabelle 11: Anzahl der übereinstimmenden und sich ändernden Diagnosen zwischen Urteil 2 und 3. Nähere Erläuterung im Text.

	B	N	S	Z	PS	ENRE D	E	
B	2 (6 %)	11	5	10	1	2		31
N	1 (50 %)	7	3	2		1		14
S			16 (80 %)	4				20
Z		2	2	3 (43 %)				7
PS					1 (50 %)	1		2
ENRE D								
E			1					1
	3	20	27	19	2	4		75

Tabelle 12: Anzahl der übereinstimmenden und sich ändernden Diagnosen zwischen Urteil 1 und 3. Nähere Erläuterung im Text.

Den Zeilen der Tabellen ist zu entnehmen, wie sich die jeweils späteren Diagnosen auf die vorhergehenden verteilen. Somit gibt jeweils die Diagonale der Tabellen die Anzahl der Übereinstimmungen zwischen den beiden psychiatrischen Urteilen bei den B und Diagnosegruppen an. Die jeweils rechte Spalte enthält die Gesamthäufigkeiten aus den früheren, die jeweils letzte Zeile die aus den späteren psychiatrischen Urteilen, verteilt über die B und Diagnosegruppen.

Von besonderem Interesse sind zunächst die Diagonalwerte, die Aufschluß über die Stabilität der Diagnosen geben. Die prozentualen Anteile der Übereinstimmungen sind in den Tabellen in Klammern angegeben. An dieser Stelle kann noch nicht zwischen Stabilität der Beurteilbarkeit der Patienten und Stabilität der Beurteilungen der Ärzte unterschieden werden. Die Stabilität der Diagnosen umfaßt hier noch beide Komponenten.

Die aufgeführten Stabilitätswerte sind unterschätzte Werte, da die Psychiater häufig mit Beschreibungen eine Diagnose umgangen haben. Die Berücksichtigung der B bei der Berechnung der Prozentwerte durch anteilige Verteilung der B auf die Diagnosegruppen erscheint jedoch auch nicht sinnvoll. In diesem Fall erfolgte eine Unterschätzung der Stabilitätswerte, da erwartet werden kann, daß bei Zwang zur Abgabe von Diagnosen weitere Diagnoseübereinstimmungen aufträten. Daß es sich hier um überschätzte Stabilitätswerte handelt, zeigt auch der Vergleich zu anderen Untersuchungen. Bei diesem Vergleich muß jedoch berücksichtigt werden, daß in diesen Untersuchungen der Zeitraum zwischen erster und zweiter Diagnose gewöhnlich größer ist als bei dieser Untersuchung (z.B. ARTHUR und GUNDERSON, 1966; ZUBIN, 1967).

Eine statistische Überprüfung des Verlaufs der Stabilitätswerte bezüglich ihrer Größe zwischen 1. und 2. (Stabilität $1/2$) und 2. und 3. (Stabilität $2/3$) psychiatrischen Urteil kann für die in den Tabellen wiedergegebenen Werte wegen der möglichen Interaktion der Meßwerträger beurteilte Patienten und beurteilende Ärzte nicht vorgenommen werden. Aus den Stabilitätswerten geht jedoch hervor, daß die Stabilität $2/3$ gegenüber der Stabilität $1/2$ vermindert ist. Diese Stabilitätsverminderung geht insbesondere zu Lasten der als neurotisch beurteilten Patienten. Bei den als schizophren und cyclothym diagnostizierten Patienten sind die Stabilitätsschwankungen minimal. Entsprechende statistische Stabilitätsprüfungen pro Arzt ergeben in keinem der 5 Fälle signifikante Unterschiede zwischen den Stabilitäten $1/2$ und $2/3$ (Binominal Test; p jeweils $>0,05$; siehe Ergebnisse unter "Differentielle Aspekte : Ärzte").

Mit diesem Ergebnis ist die Hypothese bestätigt, daß das psychologische Urteil schneller einer Stabilisierung unterliegt als das psychiatrische Urteil; das psychologische Urteil ist laut Ähnlichkeitsvergleich der Patientenfaktorenmatrizen ab 2. Urteil recht stabil. Die getroffene Aussage ist ungenau und bedarf einer Spezifizierung: Verglichen wurde der Verlauf von Patiententypen (psychologisches Urteil) und Diagnosen von Patienten (psychiatrisches Urteil). Der hier getroffene Vergleich betrifft den Inhalt von Patiententypen und Diagnosen. Davon muß der Verlauf der Differenziertheit der Patiententypen und Diagnosen unterschieden werden. Die Patiententypen aus dem psychologischen Urteil differenzieren sich auch zum 3. Urteil weiter aus (vgl. insbesondere die Ergebnisse zur Kernmatrix). Die Differenziertheit schlägt sich im psychiatrischen Urteil insbesondere in der Anzahl der B nieder. Eine hohe Anzahl von B weist auf eine undifferenzierte psychiatrische Sicht der Patienten hin. Der gleichmäßig abnehmende Anteil der B (siehe graphische Darstellung²) belegt die zunehmend differenziertere psychiatrische Sicht der Patienten. Somit trifft die schnellere Stabilisierung des psychologischen Urteils auf die Sicht der Patienten zu, d.h. die beurteilte Zugehörigkeit der Patienten zu einem Patiententyp, weniger jedoch auf das Ausmaß der Differenziertheit der Beurteilung.

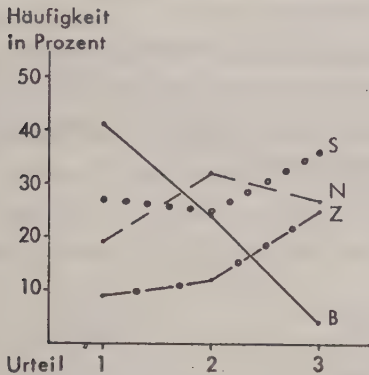
Die als schizophren diagnostizierten Patienten werden insgesamt am stabilsten beurteilt. Es folgen die mit der Diagnose N und Z belegten Patienten. Bezüglich der Stabilität $1/2$ erreichen jedoch die als neurotisch eingeschätzten Patienten den höchsten Wert.

Es ist den Tabellen zu entnehmen, daß die Verteilung der B und Diagnosen des jeweils späteren auf die B und Diagnosen des jeweils früheren Urteils von Informationsart und -ausmaß abhängig ist. B ändern sich von 1. und 2. Urteil vorrangig zu N-Diagnosen, von 2. zu 3. Urteil hauptsächlich zu Z-Diagnosen.

Vermutlich veranlaßt der Informationsinhalt aus Situation II und insbesondere der Informationsinhalt aus der sozialen Anamnese (Situation III) zur Abgabe von N-Diagnosen, die psychiatrisch besonders wichtige Information aus Situation IV und V zur Abgabe von Z-Diagnosen.

Bezüglich der Diagnoseänderungen ist insbesondere die relativ hohe Anzahl der Wandlungen von N zu S Diagnosen vom 2. zum 3. Urteil hervorzuheben. Diese Änderung kann wiederum im Zusammenhang mit der Informationsart, vermutlich weniger im Zusammenhang mit dem Informationsausmaß interpretiert werden.

Besonders deutlich schlagen sich Informationsart und -ausmaß in den Gesamthäufigkeiten der B und Diagnosen pro Urteil nieder. Die prozentualen Anteile pro Urteil von B und N-, S- und Z-Diagnosen gibt die graphische Darstellung 2 wieder.



Graph. Darst. 2: Prozentualer Anteil aller B und N-, S- und Z-Diagnosen pro Urteil

Der Anteil der B fällt erwartungsgemäß beständig. Der stärkste Anstieg der N-Diagnosen findet nach Information aus den Situationen II und III, von 1. zu 2. Urteil statt. Nach der aus den Situationen IV und V gewonnenen, psychiatrisch besonders relevanten Information fällt dann der Anteil der N-Diagnosen. Der erwartete umgekehrte Trend ist bei den Anteilen der S-Diagnosen zu beobachten. Der Anteil der Z-Diagnosen steigt in beiden Fällen, erwartungsgemäß jedoch am stärksten von 2. zu 3. Urteil.

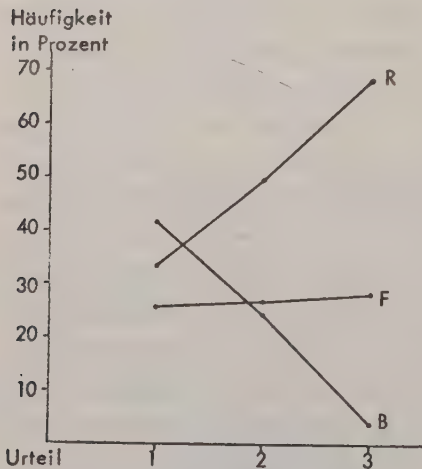
Die Angabe von Diagnosen beim 1. und 2. psychiatrischen Urteil, zu dem wenig und psychiatrisch eher irrelevante Information vorliegt, orientiert sich an den Basisraten von N, S und Z. S und N treten häufiger auf als Z. Erst im 3. Urteil, nach der psychiatrisch besonders relevanten Information aus den Situationen IV und V, findet eine Angleichung an die tatsächliche Diagnosehäufigkeit statt. Möglicherweise wird der Basisraten-Effekt noch dadurch verstärkt, daß schizophrene Erkrankungen im Verhältnis zu neurotischen und cyclothymen als besonders gefährlich bewertet werden und somit unbedingt diagnostisch erfaßt werden müssen (vgl. SCHEFF, 1963; BIERI et al., 1966). Aus einer solchen Einstellung resultiert die Tendenz, schon bei geringsten schizophrenen Anzeichen Schizophrenie zu diagnostizieren.

Die referierten Diagnoseänderungen müssen im Zusammenhang mit diesen Diagnosetendenzen gesehen werden. So ist insbesondere der Wechsel der N- zu S-Diagnosen vom 2. zum 3. Urteil abhängig vom Anstieg des S-Diagnoseanteils.

122. Validität

Zur Überprüfung der Validität der Diagnosen standen als Außenkriterium die Diagnosen der die jeweiligen Patienten behandelnden Ärzte zur Verfügung.

Der prozentuale Anteil der richtigen (R) und falschen (F) Diagnosen sowie der prozentuale Anteil der als Beschreibung (B) bewerteten Angaben sind pro Urteil in der graphischen Darstellung 3 veranschaulicht. Die Prozentwerte basieren auf den über die Merkmalsträger Ärzte und Patienten summierten Häufigkeiten.



Graph. Darst. 3: Prozentualer Anteil aller richtigen und falschen Diagnosen und aller Beschreibungen pro Urteil

Wegen der Verflechtung der beiden Merkmalsträger Ärzte und Patienten können statistische Differenz-Überprüfungen nicht vorgenommen werden. Auch bei der Interpretation der B, R und F Werte kann nicht zwischen Beurteilbarkeit der Patienten und Beurteilungsfähigkeit der Ärzte unterschieden werden. Weiterhin ist bei der Interpretation der Darstellung zu beachten, daß die B, R und F Werte voneinander abhängig sind.

Wie schon aus den Daten zur Stabilität des Urteils bekannt, steigt vom 1. zum 3. Urteil erwartungsgemäß der Anteil von Diagnosen, während der Anteil von Beschreibungen fällt. Die Veränderung der hier dargestellten Anteile von B, R und F verläuft in allen 3 Fällen sehr gleichmäßig.

Der Anstieg der R entspricht z.T. der Erwartung; die leichte Erhöhung der F widerspricht ihr.

Das Anwachsen des Informationsausmaßes sowie der psychiatrischen Relevanz der Information führt zur Erhöhung der diagnostischen Eindeutigkeit der Patienten. Entsprechend wird das psychiatrische Urteil der Ärzte sicherer, der Anteil der B fällt, der der Diagnosen steigt.

Auch durch diese Daten, die die beständige Veränderung der B, R und F belegen, wird - wie schon unter Stabilität des Urteils - die Hypothese bestätigt, daß sich das psychologische Urteil schneller verfestigt als das psychiatrische.

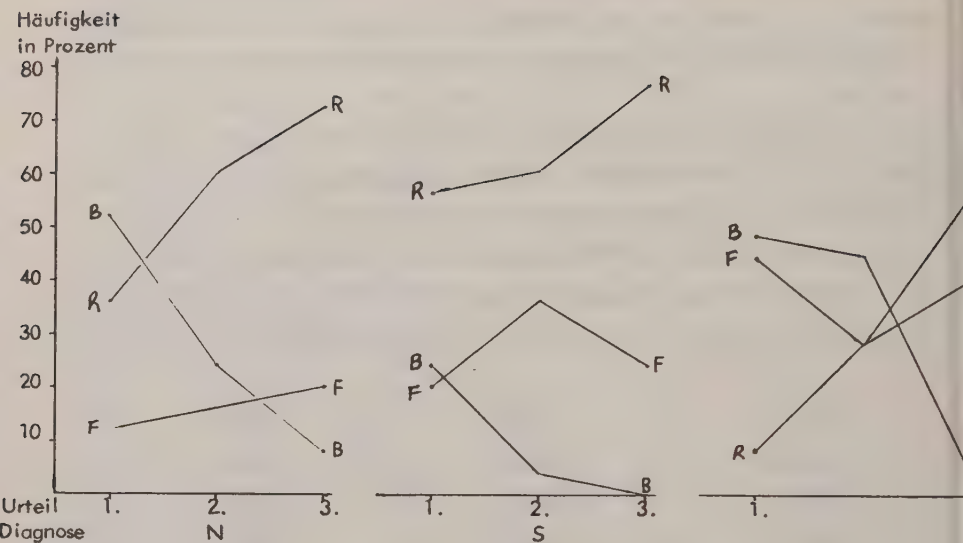
Und anhand des Anteils der inhaltlich angegebenen F wird der schon unter Stabilität des psychiatrischen Urteils aufgezeigte Sachverhalt bekräftigt, daß vermutlich der Informationsinhalt Einfluß auf das psychiatrische Urteil nimmt (Tabelle 13).

Urteil 1	Urteil 2	Urteil 3
31% N	45% N	9% N
31% S	20% S	38% S
26% Z	10% Z	24% Z
11% Psychopath	15% Psychopath	9% Psychopath
	10% Entwicklungskrise	19% endo-reaktive Depression

Tabelle 13: Prozentualer Anteil von allen falschen Diagnosen pro Urteil

Der prozentuale Anteil der Falschdiagnosen N ist nach der Information aus den Situationen II und III und der prozentuale Anteil der Falschdiagnosen S nach der Information aus den Situationen IV und V sowohl im Verhältnis zu den Anteilen der anderen F pro jeweiligem Urteil als auch im Verhältnis zu den eigenen F in den anderen Urteilen am höchsten.

Deutlich schlägt sich dieser vermutete Informationsinhalt-Effekt auch im prozentualen Anteil der R pro Diagnosegruppe N, S und Z nieder (graphische Darstellung 4).



Graph. Darst. 4: Prozentualer Anteil aller richtigen und falschen Diagnosen und Beschreibungen pro Diagnosegruppe und Urteil

Der Anteil der R steigt bei der Diagnosegruppe der N am stärksten von Urteil 1 nach 2, bei der Diagnosegruppe der S dagegen von Urteil 2 nach 3. Ein erwartungsgemäßer, wenn auch nicht so deutlicher Effekt ist bei der Diagnosegruppe der Z festzustellen.

Die Anteile von R, B und F stehen pro Urteil in gegenseitiger Abhängigkeit. Der Anstieg des Anteils richtiger Diagnosen von 1. bis 3. Urteil hat daher eine Veränderung der Anteile von B und F zur Folge.

Die Rangkorrelationen zwischen R und B und R und F über die Patienten, wobei die über die Ärzte summierten Anzahlen von R, F und B pro Urteil und Patient als Meßwerte fungieren, zeigen, daß sowohl beim ersten als auch noch beim 2. Urteil die B signifikant Einfluß auf die R nehmen.

Patienten, die beschrieben werden, werden nicht richtig beurteilt und umgekehrt. Erst ab zweitem Urteil - nachdem sich die Anzahl der B vermindert hat - tritt auch der erwartete signifikante Zusammenhang zwischen R und F auf (Tabelle 14).

Urteil	1	2	3
R/B	-.766	-.635	-.257
R/F	-.413	-.675	-.958
B/F	-.253	-.082	-.020

Tabelle 14: Zusammenhang zwischen richtigen, falschen Diagnosen und Beschreibungen pro Urteil (Rangkorrelationskoeffizienten)
 $r_s \geq .43$, $p \leq 0,05$; $r_s \geq .63$, $p \leq 0,01$

Bei einer Interpretation des Verlaufs der Anteile von B, R und F müssen diese Zusammenhänge berücksichtigt werden.

Wie aus der obigen graphischen Darstellung hervorgeht, ist die Beziehung zwischen den Anteilen von F und B für den Durchschnitt der von den 5 Ärzten beurteilten Patienten der Diagnosegruppen S und Z von besonderer Art:

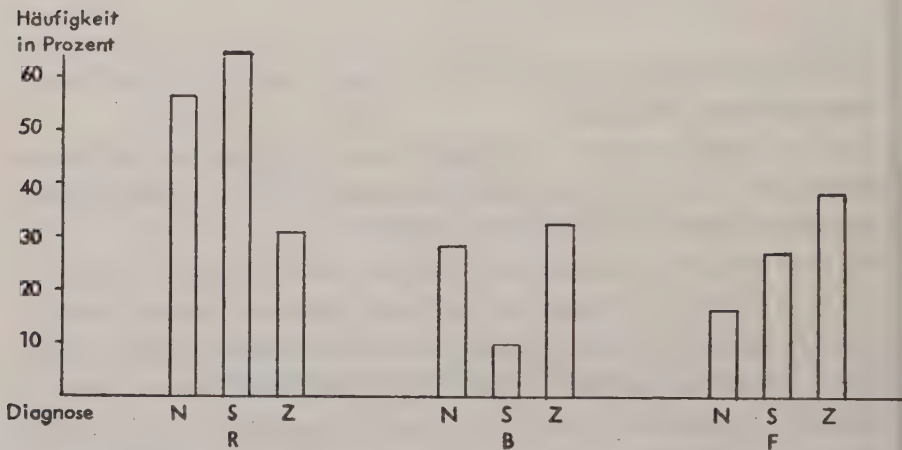
Mit einem starken Abfall des Anteils der B geht ein beträchtlicher Anstieg des Anteils von F und mit einer geringfügigen Verminderung des Anteils der B auch ein leichter Abfall des Anteils der F einher. Danach ist ein schneller Übergang von B zu Diagnosen bei den beiden Diagnosegruppen z.T. voreilig, d.h. nicht nur der Anteil der R, sondern auch der der F erhöht sich. Dieser Effekt tritt bei der Gruppe der N nicht auf. Bei dieser Gruppe verläuft die Veränderung des Anteils von B jedoch auch relativ gleichmäßig.

B sind Ausdruck für die undifferenzierte Sicht der Patienten, damit aber auch gleichzeitig Ausdruck für die diagnostische Mehrdeutigkeit der Patienten, bzw. - bei den hier über Patienten und Ärzte zusammengefaßten Werten - für die Urteilsunsicherheit der diagnostizierenden Ärzte. Demnach trägt die Information aus Situation II und III insbesondere bei der Gruppe der S zur Beurteilbarkeit der Patienten und der subjektiven Sicherheit der diagnostizierenden Ärzte bei.

Entsprechend steigt die Anzahl der Diagnosen. Im Gegensatz zur Gruppe der S findet der entscheidende Abfall des Anteils von B bei der Gruppe der Z erst nach der Information aus Situation IV und V, also von Urteil 2 auf 3 statt. Zur Z-Diagnose bedarf es also mehr, insbesondere psychiatrisch relevanter Information: Die Gruppe der Z wirkt somit, vor allem im Verhältnis zur Gruppe der S bis zu Situation III diagnostisch mehrdeutiger und löst bei den Ärzten Urteilsunsicherheit aus.

Bei der Gruppe der N tritt ein entsprechender Effekt nur in sehr geringem Ausmaß auf.

Aufgrund der genannten Ergebnisse erscheinen die Patienten der Gruppe Z besonders schwer zu diagnostizieren. Das zeigt sich sowohl am Verhältnis der Anteile von R, F und B für alle drei Urteile insgesamt (graphische Darstellung 5) als auch pro Urteil (siehe graphische Darstellung 4).



Graph. Darst. 5: Prozentualer Anteil aller richtigen Diagnosen, Beschreibungen und falschen Diagnosen pro Diagnosegruppe

Patienten der Gruppe S sind im Verhältnis zu denen der Gruppe N, insbesondere aber zu denen der Gruppe Z diagnostisch eindeutiger und erhalten im Durchschnitt die größte Anzahl richtiger Diagnosen. Besonders deutlich ist dieser Sachverhalt beim 1. Urteil zu beobachten. Bei diesem Urteil werden die Patienten der Gruppe Z nur selten richtig, aber häufig falsch diagnostiziert oder beschrieben.

Die F pro Patientengruppe geben Aufschluß darüber, welche differential-diagnostischen Schwierigkeiten bei der jeweiligen Gruppe aufkamen (Tabelle 15).

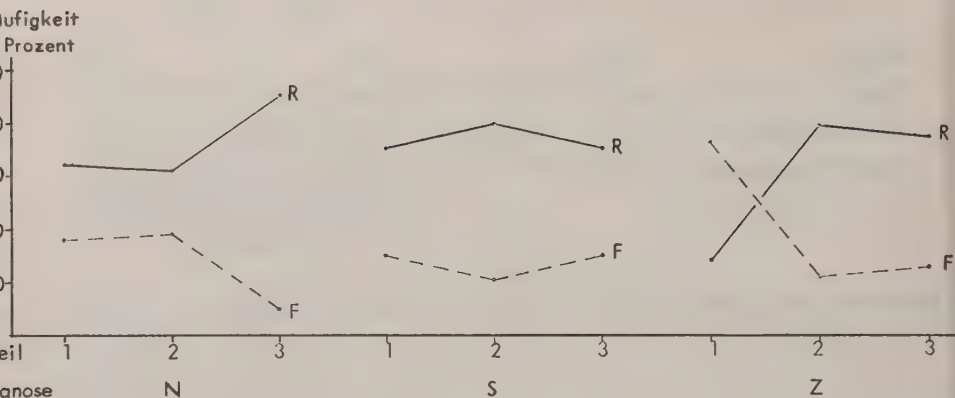
	N		S		Z
33%	Z	55%	N	54%	S
25%	S	35%	Z	21%	N
25%	Psychopath	10%	Entwicklungs- krise	14%	Psychopath
17%	endo-reakt. Depression			7%	endo-reakt. Depression
				4%	endogene Depression bei Maniker

Tabelle 15: Prozentualer Anteil von allen falschen Diagnosen pro Diagnosegruppe N, S und Z

Die Verteilung der Anteile von F auf die Gruppe der N ist im Verhältnis zur Gruppe der S und insbesondere zu der der Z recht gleichmäßig. Differential-diagnostische Schwierigkeiten treten damit gegenüber den beiden Gruppen S und Z und der Gruppe Psychopathie etwa gleich stark auf.

Dagegen sind schizophrene Patienten gegenüber neurotischen diagnostisch schwieriger abgrenzbar als gegenüber cyclothymen Patienten, und bei cyclothymen Patienten ergibt sich die differential-diagnostische Schwierigkeit insbesondere gegenüber schizophrenen Patienten.

Das Ergebnis, die Gruppe der Z sei im Verhältnis zur Gruppe der N und insbesondere zu der Gruppe S schwerer zu diagnostizieren, muß aufgrund der Tatsache, daß sich die diagnostizierenden Ärzte bei der Abgabe der Diagnosen auch an den Basisraten dieser Diagnosen orientierten (siehe Ergebnisse unter Stabilität des Urteils), relativiert werden. Der Verlauf von R und F über die drei Urteile pro Diagnosegruppe sowie das Verhältnis von jeweils R und F zwischen den Diagnosegruppen ändert sich beträchtlich, wenn die Anzahl von R pro Diagnosegruppe und Urteil in Beziehung gesetzt wird zur jeweiligen Gesamtzahl der entsprechenden abgegebenen Diagnosen (graphische Darstellung 6). Die Darstellung enthält auch die komplementäre F-Diagnose.



Graph. Darst. 6: Prozentualer Anteil richtiger und falscher Diagnosen an den pro Diagnosegruppe und Urteil jeweils insgesamt abgegebenen Diagnosen

Diese Kurven können wie bisher interpretiert werden.

Bei der Interpretation der Kurven ist jedoch - insbesondere im Hinblick auf die Höhe der R - zu berücksichtigen, daß sich durch die Bezugnahme auf die jeweilige Gesamtzahl der entsprechenden abgegebenen Diagnosen ein völlig unrealistisches Bild ergibt, denn die Ärzte können sich beim individuellen diagnostischen Fall mit ihrer Diagnose nicht an Basisraten ausrichten. Aufschlußreicher ist dagegen der Vergleich dieser Kurven mit den Kurven aus der graphischen Darstellung 2, die den Verlauf der Anteile abgegebener N-, S- und Z-Diagnosen pro Urteil enthält. So besteht bei den Diagnosegruppen N und S ein deutlicher, bei der Diagnosegruppe Z ein weniger deutlicher negativer Zusammenhang zwischen dem Verlauf der abgegebenen und richtigen Diagnosen bzw. ein positiver Zusammenhang zwischen dem Verlauf der abgegebenen und falschen Diagnosen. Dieses Ergebnis stimmt überein mit dem schon referierten, daß eine (voreilige) Erhöhung der Anzahl von Diagnosen einhergeht mit der Zunahme falscher Diagnosen. Dieses Ergebnis kann nicht ohne weiteres generalisiert werden. In der Untersuchung fühlten sich die Ärzte selbst bei unzureichender Information zur Angabe von Diagnosen gezwungen, wenn sie auch - vor allem beim 1. Urteil - manchmal einer Diagnoseangabe durch Beschreibungen auswichen. In der Praxis werden Diagnosen gewöhnlich erst nach weitaus mehr Information formuliert.

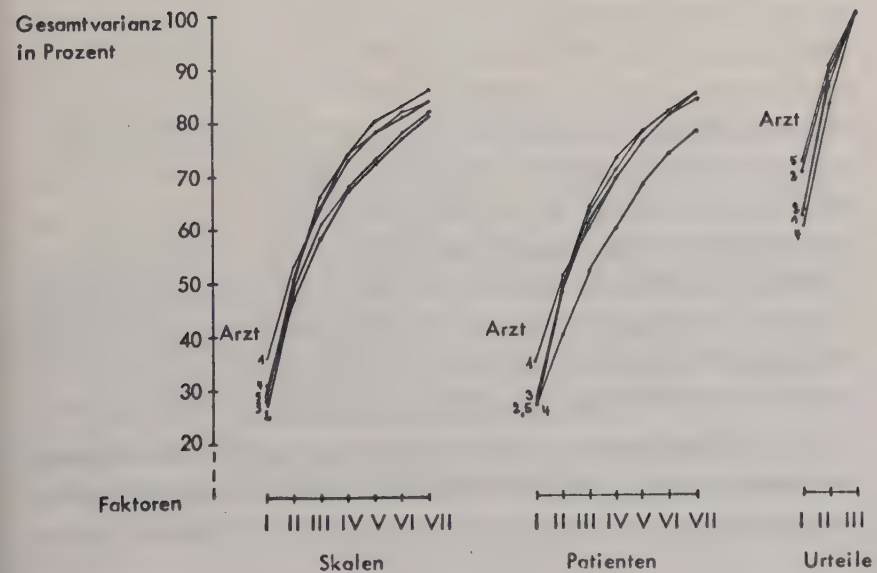
2. Differentielle Aspekte

21. Psychologisches Urteil

Pro Arzt wurde eine 3 Modi-FA durchgeführt. Die Modi sind Skalen, beurteilte Patienten und die Beurteilungswiederholungen.

211. Differenziertheit der Urteile

Die folgende graphische Darstellung 7 veranschaulicht die kumulierten Varianzanteile der ersten 7 unrotierten Skalen-, der ersten 7 unrotierten Patienten- und der drei unrotierten Beurteilungswiederholungsfaktoren pro Arzt.



Graph. Darst. 7: Kumulierte prozentuale Gesamtvarianzanteile der unrotierten Faktoren pro Modus und Arzt

Bei Arzt 1, 3 und 4 können jeweils 4 Skalenfaktoren als bedeutsam angesehen werden, bei Arzt 2 und 5 scheinen es nur drei zu sein. Arzt 1 fällt durch einen verhältnismäßig varianzstarken 1. Skalenfaktor auf. Die Ärzte unterscheiden sich damit im Ausmaß der Differenziertheit ihrer impliziten Persönlichkeitstheorien bzw. kognitiven Beurteilungsräume. Die Ärzte 2 und 5 verfügen über eine geringere Anzahl von Beurteilungsdimensionen, und Arzt 1 tendiert dazu, eine relativ große Anzahl von Beurteilungsskalen als gleichbedeutend zu verwenden. Damit wird belegt, daß sich die individuellen impliziten Persönlichkeitstheorien auf der Ebene struktureller Aspekte erheblich unterscheiden.

Mit Ausnahme von Arzt 4 sind die entsprechenden strukturellen Differenzen bei den Patientenfaktoren weniger gravierend. Bei den Ärzten 2 bis 5 können jeweils 3 Faktoren als bedeutsam gelten. Lediglich Arzt 1 weicht von den anderen Ärzten durch ein verhältnismäßig globales Urteil über die Patienten ab. Er unterscheidet nur 2 Patiententypen, wobei schon der erste eine relativ große Anzahl von Patienten auf sich vereinigt. Und Arzt 4 sieht die Patiententypen relativ "spezifisch". Hervorzuheben ist, daß bei diesen individuellen Urteilsanalysen nur maximal 3 Patiententypen auftreten. Bei den bisherigen Analysen, bei denen alle 5 Ärzte in die Berechnung eingingen, kamen ab 2. Urteil 4 Patiententypen zustande. Diese Differenz ist ein Hinweis darauf, daß sich die im Urteil implizierten Patiententypen zwischen den Ärzten inhaltlich erheblich unterscheiden müssen; bei einer Gesamtberechnung kristallisiert sich folglich eine größere Anzahl von Patiententypen heraus.

Von den Beurteilungswiederholungsfaktoren ist jeweils nur der 1. Faktor bedeutsam. Er erfaßt im Durchschnitt 65,2% der Gesamtvarianz, womit angezeigt wird, daß die 3 Urteile im Durchschnitt relativ ähnlich sind. Dieses Ergebnis stimmt mit den bisherigen Analysen überein. Urteil 2 und 3 erwiesen sich bezüglich aller 3 Modi als recht ähnlich, und zwischen Urteil 1 und 2 bestand auch noch gewisse Übereinstimmung.

Es treten deutliche Unterschiede in der Varianzstärke des jeweils 1. Beurteilungswiederholungsfaktors zwischen den Ärzten auf. Diese Unterschiede sollten jedoch nicht überbewertet werden, da dieser erste Faktor jeweils schon sehr varianzstark ist.

212 Skalen

Es wurden pro Arzt 4 Faktoren extrahiert, zur Interpretation varimaxrotiert und die 5 Faktorenmatrizen auf ihre Ähnlichkeit hin überprüft (Tabelle 16).

Arzt				
1	.715	.835	.832	.867
2		.702	.696	.697
3			.850	.846
4				.853

Tabelle 16: Ähnlichkeitskoeffizienten aus dem Vergleich der 5 Skalenfaktorenmatrizen

Die Faktorenstrukturähnlichkeit zwischen den Ärzten 1, 3, 4 und 5 ist nicht hoch, es bestehen aber noch deutliche Verwandtschaftsverhältnisse. Dagegen weicht die Faktorenstruktur des Arztes 2 schon erheblich von denen der anderen Ärzte ab. Dieser Sachverhalt ist möglicherweise z.T. darauf zurückzuführen, daß bei diesem Arzt ebenfalls 4 Faktoren zum Faktorenvergleich extrahiert wurden, obwohl nur 3 Faktoren eindeutig bedeutsam sind. Dieses Vorgehen dürfte jedoch nicht ausschlaggebend für die Verminderung der Ähnlichkeiten sein, denn bei Arzt 5 können auch nur 3 Faktoren als bedeutsam gelten. Bei diesem Arzt tritt ein entsprechender Effekt aber nicht auf, so daß bei Arzt 2 ein inhaltlich beträchtlich andersartiger kognitiver Beurteilungsraum vermutet werden kann.

Damit, aber auch aufgrund der anderen Ähnlichkeitskoeffizienten ist die Hypothese bestärkt, daß sich die individuellen impliziten Persönlichkeitstheorien inhaltlich voneinander unterscheiden. Jedenfalls sind diese individuellen Unterschiede größer als diejenigen, die bei zunehmender Information über die Patienten zwischen den jeweils durchschnittlichen impliziten Persönlichkeitstheorien festgestellt wurden. Damit ist die implizite Persönlichkeitstheorie gegenüber zunehmender Information über dieselben Personen, aber auch gegenüber unterschiedlich zu beurteilenden Personengruppen relativ invariant (vgl. z.B. NORMAN und GOLDBERG, 1966).

Dagegen treten z.T. beträchtliche interindividuelle Variationen auf.

Auf die Interpretation der individuellen Skalenfaktoren wird aus Raumgründen verzichtet.

213. Patienten

Es wurden pro Arzt 3 Faktoren extrahiert, die Faktoren zur Interpretation wie üblich varimaxrotiert und die 5 Faktorenmatrizen verglichen (Tabelle 17).

Urteil	2	3	4	5
1	.635	.699	.396	.764
2		.629	.710	.665
3			.553	.856
4				.570

Tabelle 17: Ähnlichkeitskoeffizienten aus dem Vergleich der 5 Patientenfaktorenmatrizen

Die mittlere Faktorenstrukturähnlichkeit ist gering (.647), die Streuung der Koeffizienten relativ groß (.125). Es bestehen z.T. keine Faktorenstrukturähnlichkeiten, z.B. zwischen Arzt 1 und 4, aber auch 3 und 4. Eine eindeutige Strukturähnlichkeit weisen lediglich die Faktorenmatrizen von Arzt 3 und 5 auf. Die in den Urteilen der Ärzte implizierten Patiententypen sind damit im Durchschnitt inhaltlich nur sehr wenig ähnlich.

Dieses Ergebnis steht in einer interessanten Beziehung zu dem entsprechenden Unterskalen-Modus: Mit Hilfe interindividuell immerhin noch relativ ähnlichen Beurteilungsdimensionen werden die Patienten interindividuell z.T. völlig unterschiedlich beurteilt, so daß es zu unähnlichen Patiententypen kommt. D.h., die Beurteilungsskalen werden zwar ähnlich interpretiert, die Patienten anhand dieser Skalen jedoch beträchtlich unterschiedlich eingeschätzt.

Wie schon ausgeführt, können die Patientenkommunalitäten als Ausdruck für die Urteils-Bestimmtheit der Patienten interpretiert werden. Bei keinem Arzt treten signifikante Differenzen zwischen den Diagnosegruppen bezüglich der Höhe der Kommunalitäten auf. (Kruskal-Wallis Test; $H_1 = 1,50$; $H_2 = 3,54$; $H_3 = 3,66$; $H_4 = 4,16$; $H_5 = 0,38$ bei jeweils $df = 2$; p jeweils $> 0,05$).

Auch die interindividuellen Unterschiede pro Diagnosegruppe erreichen keine Signifikanz (Friedman Test; $\chi^2_N = 6,2$; $\chi^2_S = 2,4$; $\chi^2_Z = 4,0$ bei jeweils $df = 4$; p jeweils $> 0,05$). Das Ausmaß der Urteilsbestimmtheit der 3 Diagnosegruppen ist somit unter dem Gesichtspunkt intraindividuelle wie interindividuelle Differenzen nicht unterschiedlich.

Die pro Arzt errechneten mittleren Q-Korrelationen zwischen jeweils beurteilten realen Patienten und den 9 typischen, aus der Vorstellung heraus beurteilten Patienten waren als Maß für die Tendenz eines Arztes interpretiert worden, sich beim Urteil

über reale Patienten an einer Vielzahl von Patientenstereotypen zu orientieren. Von Interesse sind zunächst diesbezügliche intraindividuelle Unterschiede zwischen den drei Urteilen, d.h. die Abhängigkeit des Ausmaßes der differenzierten Urteilsstereotypie von Inhalt und Ausmaß der Information über die Patienten. Dabei kann wiederum von der Hypothese ausgegangen werden, daß zunehmende Information zu einer Erhöhung des beurteilten stereotypen Symptomreichtums führt, da die Patienten zunehmend durch den Vergleich mit mehreren Patiententypen beurteilt werden. Bei keinem der Ärzte konnte die Hypothese bestätigt werden (Friedman Test; $\chi^2_1 = 2,3$; $\chi^2_2 = 1,5$; $\chi^2_3 = 0,00$; $\chi^2_4 = 0,7$; $\chi^2_5 = 0,3$ bei jeweils $df = 2$; p jeweils $> 0,05$). Auch die pro Arzt überprüften Unterschiede zwischen dem Ausmaß des Symptomreichtums der Diagnosegruppen N, S und Z bei Urteil 1, 2 und 3 erwiesen sich lediglich in einem der 15 Fälle als signifikant (Kruskal-Wallis Test; Tabelle 18); ein signifikanter Wert ist jedoch schon beinahe nach Zufall zu erwarten. Somit hat pro Arzt weder Informationsausmaß und -art noch die jeweiligen Patientengruppen Einfluß auf das Ausmaß der Orientierung an der Symptomatik typischer Patienten.

Arzt	Urteil		
	1	2	3
1	1,14	1,94	3,44
2	0,72	0,86	1,22
3	0,94	1,34	1,52
4	5,82	1,46	5,30
5	3,86	4,82	3,62

Tabelle 18: H-Werte; Erläuterung im Text. Gruppengröße jeweils $N = 5$;
 $H = 5,78$, $p = 0,049$; $H = 7,98$, $p = 0,010$

Interindividuelle Unterschiede zwischen den Ärzten bezüglich dieser stereotypen Symptomvielfalt geben Aufschluß darüber, wie weit sich die Ärzte bei der Beurteilung der Patienten in unterschiedlichem Ausmaß an ihrer Vorstellung von typischen Patienten orientieren oder die Individualität eines jeden Patienten im Urteil beachten. Die Differenzen sind in allen 3 Urteilen signifikant (Friedman Test; $\chi^2_1 = 9,91$; $p < 0,05$; $\chi^2_2 = 15,76$; $p < 0,01$; $\chi^2_3 = 10,12$; $p < 0,05$ bei jeweils $df = 4$).

Damit steht fest, daß sich die Ärzte bezüglich der differenzierten stereotypen Beurteilung signifikant unterscheiden.

In bisherige Überprüfungen gingen nur Werte ein, die das Ausmaß der differenzierten stereotypen Beurteilung widerspiegeln. Ärzte, die hohe Werte in dieser Variablen errechnen, urteilen im Sinne ihrer Stereotype, die sie von typischen Patienten haben, differenziert, berücksichtigen jedoch nicht ausreichend die Individualität der Patienten. Es ist nun von Interesse, in welchem Ausmaß jeder Arzt bei der Beurteilung der Patienten einer Diagnosegruppe sein Urteil über diese Patienten an einem typischen, dieser Diagnosegruppe zuordenbaren Patienten ausrichtet. Im Sinne dieser Frage kann pro Arzt dann überprüft werden, ob sich die Höhe der Q-Korrelationen zwischen den neurotischen Patienten und jeweils dem typischen hysterischen, hypochondrischen und depressiven neurotischen Patienten zwischen den Urteilen unterscheidet. Entsprechende intraindividuelle Überprüfungen können auch für die Gruppe der Schizophrenen bzw. Cyclothymen vorgenommen werden, wobei jeweils Korrelationen zu den typischen 3 schizophrenen bzw. 2 cyclothymen Patienten in die Berechnung eingehen. Bei diesen insgesamt 40 Überprüfungen traten nur 6 signifikante Differenzen auf (Friedman Test; $df = 2$, p sechsmal $< 0,05$). Es erscheint unrentabel, darauf einzugehen, bei welchen Patientengruppen und typischen Patienten diese Differenzen mit den entsprechenden χ^2 -Werten vorkamen. Ausmaß und Art der Information über die Patienten scheinen nur wenig Einfluß auf das Ausmaß der stereotypen Beurteilung pro Diagnosegruppe zu nehmen. Dagegen ist die Abhängigkeit der stereotypen Beurteilung des individuellen Patienten von Informationsart und -ausmaß häufig eindeutig gegeben, wie steigende Korrelationen von .23, .59 bis .76 z.B. bei Arzt 5 belegen.

Entsprechend treten signifikante interindividuelle Differenzen zwischen den Ärzten bezüglich der stereotypen Beurteilung insbesondere nur dann gehäuft auf, wenn Korrelationen zwischen Beurteilungen eines individuellen und eines typischen, derselben Krankengruppe zugehörigen Patienten gegenübergestellt werden. Die Überprüfung pro Patientengruppe erbrachte daher nur – bei jeweils 8 Differenzen – im 1. Urteil eine signifikante, im 2. Urteil 3 hoch signifikante und im 3. Urteil eine hoch signifikante und eine signifikante Differenz zwischen den Ärzten. Die Anzahl der signifikanten Differenzen ist bei diesem Überprüfungsansatz relativ gering (Friedman Test; Tabelle 19).

	Urteil	1	2	3
Neurotiker	depress.	0.32	14.88	11.84
	hypoch.	4.48	4.96	7.92
	hyst.	7.52	13.92	13.92
Schizophrene	1. Schub	7.52	5.76	5.60
	Defektschizo.	4.52	3.20	6.40
	Hebephre.	2.40	3.84	7.20
Cyclothyme	man. Phase	10.72	3.68	5.60
	endog. Depress.	5.44	13.92	9.12

Tabelle 19: χ^2 -Werte; Erläuterung im Text. df jeweils = 4;
 $\chi^2 \geq 9.49$, $p \leq 0.05$; $\chi^2 \geq 13.28$, $p \leq 0.01$

Überprüfungen, wie weit Informationsart und -ausmaß den Grad der stereotypen Beurteilung bestimmen, sowie Prüfungen interindividueller Stereotypietendenzen im Urteil sind somit nur dann ausreichend aufschlußreich, wenn Stereotypietendenzen beim jeweiligen individuellen Patienten erfaßt werden. Die Mitglieder einer Patientengruppe erscheinen - auch im 3. Urteil - so heterogen, daß die Ermittlung von Stereotypietendenzen bei der gesamten Gruppe und folgende intra- und interindividuelle Unterschiedsüberprüfungen wenig sinnvoll sind.

214. Urteile

Um pro Arzt überprüfen zu können, welches der 3 Urteile sich am stärksten von den beiden anderen unterscheidet, wurden bei diesem Modus jeweils 2 Faktoren extrahiert, obwohl nur eindeutig ein Faktor bedeutsam ist. Die Extraktion von 2 Faktoren bei nur 3 Variablen (3 fache Beurteilungswiederholung) ist beim deskriptiven Hauptkomponentenmodell gestattet, da laut diesem Modell die Anzahl der zu extrahierenden Faktoren lediglich durch die Anzahl der Variablen begrenzt wird.

Die Faktoren wurden - wie bisher - zur Interpretation varimaxrotiert und die Ähnlichkeit der 5 Faktorenmatrizen ermittelt.

Auf die Darstellung der Ähnlichkeitskoeffizienten kann verzichtet werden. Der Durch-

schnitt der Koeffizienten ist hoch (.939), die Streuung gering (.045). Die Höhe der Koeffizienten kann sicherlich z.T. auf die geringe Anzahl von Variablen (3 Urteile) zurückgeführt werden.

Anhand der Ergebnisse der pro Urteil durchgeführten Analysen, in die jeweils alle Ärzte eingingen, war eine deutliche Veränderung von 1. zu 2. Urteil und ab 2. Urteil eine Stabilisierung nachgewiesen worden. Die Interpretation der hier errechneten Faktoren liefert nun einen interessanten Aufschluß (Tabelle 20).

		Arzt	1	5	2	3	4
Faktor	Urteil						
I	1		.770	-.804	-.961	.978	-.999
	2		.638	-.594	-.211	.152	.000
	3		.015	-.024	.181	-.145	-.011
II	1		.167	.159	.022	-.015	.000
	2		-.178	-.179	.590	-.639	.722
	3		-.970	-.971	.807	-.769	-.692

Tabelle 20: Ladungen der 3 Urteile auf den beiden rotierten Faktoren pro Arzt

Bei den Ärzten 1 und 5 differiert Urteil 1 am stärksten von Urteil 3. Urteil 2 steht in engerer Beziehung zu dem durch Urteil 1 als zu dem durch Urteil 3 aufgemachten Faktor; Urteil 2 ist Urteil 1 ähnlicher als Urteil 3.

Dagegen findet bei den Ärzten 2, 3 und 4 die schon unter "Allgemeinen Gesichtspunkten" aufgezeigte stärkste Urteilsänderung von Urteil 1 zu Urteil 2 statt. Bei den Ärzten 2 und 3 teilt Urteil 2 mehr Varianz mit dem durch das 3. Urteil als mit dem durch das 1. Urteil bestimmten Faktor; Urteil 2 ist also Urteil 3 ähnlicher als Urteil 1 und Urteil 1 und 3 differieren wie bei den Ärzten 1 und 5 am meisten. Im Gegensatz dazu sind bei Arzt 4 Urteil 1 und 2 am unterschiedlichsten. Urteil 3 liegt auf der Ähnlichkeitsdimension zwischen Urteil 1 und 2, ist jedoch Urteil 2 eindeutig ähnlicher.

Damit ergeben sich zwei Arzttypen. Bei den Ärzten 2, 3 und 4 erfolgt die stärkste Urteilsänderung von Urteil 1 nach 2, dagegen bei den Ärzten 1 und 5 von Urteil 2 nach 3.

Mit dieser interindividuellen Betrachtung wird das Ergebnis unter "Allgemeine Gesichtspunkte" relativiert. Es wird deutlich, daß eine alle Urteiler umfassende Analyse zu grobe Ergebnisse liefert.

Von Urteil 1 bis 3 sich ändernde Entscheidungsschwierigkeit beim Einstufen von Patienten anhand der Skalen des Polaritäten-Profiles scheint die Moderator-Variable zu sein, aufgrund derer die Trennung in die beiden Arzttypen erfolgt. Bei den Ärzten 1 und 5 fällt - laut Skala "Entscheidung: sehr leicht vs. sehr schwer" - die subjektive Entscheidungsschwierigkeit von Urteil 1 bis 3 sehr hoch signifikant ab (Friedman Test; $\chi^2_1 = 15,63$, $p < 0,001$; $\chi^2_5 = 24,03$, $p < 0,001$ bei jeweils $df = 2$). Bei den Ärzten 2 und 4 ergibt sich keine signifikante Veränderung der subjektiven Entscheidungsschwierigkeit (Friedman Test; $\chi^2_2 = 0,93$, $p > 0,05$; $\chi^2_4 = 2,63$, $p > 0,05$ bei jeweils $df = 2$), während bei Arzt 3 eine signifikante Verminderung der Entscheidungsschwierigkeit auftritt (Friedman Test; $\chi^2_3 = 7,43$ bei $df = 2$, $p < 0,05$).

Für die folgenden Überlegungen ist nun weiterhin noch wichtig, daß die interindividuellen Differenzen bezüglich des Ausmaßes an Entscheidungsschwierigkeit im 1. und 2. Urteil sehr hoch, im 3. Urteil aber nicht signifikant sind (Friedman Test; $\chi^2_1 = 25,57$, $p < 0,001$; $\chi^2_2 = 21,57$, $p < 0,001$; $\chi^2_3 = 6,97$, $p > 0,05$ bei jeweils $df = 4$) und daß die signifikanten Differenzen im 1. und 2. Urteil jeweils durch eine erhöhte Entscheidungsschwierigkeit bei den Ärzten 1, 3 und 5 im Verhältnis zu den Ärzten 2 und 4 hervorgerufen werden. (Wilcoxon und Wilcox; aufgeführt sind signifikante Differenzen). 1. Urteil: ($p_{2/1} < 0,01$; $p_{4/1} < 0,05$; $p_{2/5} < 0,01$; $p_{4/5} < 0,05$) 2. Urteil: ($p_{2/1} < 0,01$; $p_{4/1} < 0,01$).

Läßt man Arzt 3, bei dem die intraindividuelle Differenz zwischen den drei Urteilen lediglich auf dem 5%-Niveau signifikant ist, unberücksichtigt, dann kann folgende Aussage getroffen werden: Anfängliche, erhebliche Entscheidungsschwierigkeiten verzögern eine starke Veränderung des Urteils. (Arzt 1 und 5). Dagegen führt ein gleichbleibendes geringes Maß an Entscheidungsschwierigkeit zu einer relativ frühen Urteilsstabilisierung. (Arzt 2 und 4). Dieses Ergebnis ist unmittelbar plausibel. Von Interesse ist es nun festzustellen, worauf diese Unterschiede zurückzuführen sind. Am ehesten sind die genannten Ergebnisse noch im Rahmen der Untersuchungen zur kognitiven Komplexität interpretierbar (BIERI et al., 1966). Demnach scheinen die Ärzte 2 und 4 im Verhältnis zu den Ärzten 1 und 5 kognitiv komplexer zu sein. Ihr psychologisches Urteil erreicht nach der psychologisch relevanten Information aus Situation III eine gewisse

Stabilität, während bei den Ärzten 1 und 5 erst die psychiatrisch besonders wichtige Information aus den Situationen IV und V nach langer Unsicherheit zur entscheidenden Veränderung des Urteils beiträgt. Das letzte Urteil scheint für diese Ärzte etwa im Sinne eines Recency-Effekts das wichtigste zu sein; sie werfen ihr bisheriges Urteil trotz der psychologischen Information aus Situation III gewissermaßen um.

Die aus dem Begriff der kognitiven Komplexität ableitbare Erwartung, daß nämlich die Ärzte 1 und 5 ihr 3. Urteil auch entsprechend stärker an stereotyper Symptomvielfalt ausrichten, kann jedoch nicht bestätigt werden. Überprüft wurde diese Hypothese durch multiplen Vergleich der 4 Ärzte nach Wilcoxon und Wilcoxon. Statistisch war zu erwarten, daß a) die Differenz im Ausmaß der stereotypen Symptomvielfalt jeweils zwischen den Ärzten 1 und 5 und den Ärzten 2 und 4 nicht signifikant ist, b) die Ärzte 1 und 5 sich jedoch jeweils von den Ärzten 2 und 4 durch ein größeres Ausmaß an beurteiltem stereotypen Symptombreichtum signifikant unterscheiden.

Die Differenzen zwischen den Ärzten 1 und 5 sind entsprechend der Erwartung nicht signifikant ($p > 0,05$), jedoch erweist sich die Differenz zwischen den Ärzten 2 und 4 als signifikant (zweiseitig; $p < 0,05$). Die einzige weitere, der Erwartung entsprechende signifikante Differenz tritt zwischen Arzt 4 und 5 auf ($p < 0,05$).

Die gegebene Interpretation der Ergebnisse im Rahmen der Untersuchungen zur kognitiven Komplexität besitzt zwar eine gewisse theoretische Bedeutung, aus praktisch-psychiatrischer Sicht erscheinen die Ergebnisse jedoch in einem anderen Licht: Ärzte (hier die Ärzte 2 und 4), denen das psychologische Beurteilen der Patienten anhand der Skalen laut eigenen Angaben wenig Schwierigkeiten bereitet, bilden sehr schnell ein stabiles psychologisches Urteil über die Patienten; zur psychiatrischen Diagnose wichtige Information beeinflusst das Urteil nur noch wenig. Diese Ärzte trennen somit relativ eindeutig zwischen psychologischem und psychiatrischem Urteil. In den Situationen IV und V werden die die psychiatrischen Klassen bestimmenden Symptome abgefragt, die ermittelten Symptome beeinflussen nur noch wenig das psychologische Bild der Patienten und wahrscheinlich ist die psychologische Betrachtung der Patienten von geringer Bedeutung für die psychiatrische Diagnose. Dieses Vorgehen kommt dem typisch psychiatrischen Vorgehen relativ nahe und entsprechend haben die Ärzte 2 und 4 im Verhältnis zu den Ärzten 1 und 5 die längere psychiatrische Erfahrung (durchschnittlich 5,25 zu 3,00 Jahren). Dagegen scheinen die Ärzte 1 und 5 den Schwerpunkt ihrer Beurteilung auf eine psychologische Beurteilung der Patienten gelegt zu haben. Demgemäß

berücksichtigen sie bei ihrer psychologischen Beurteilung die psychiatrisch besonders wichtige Information aus den Situationen IV und V.

Arzt 3, der von allen Ärzten der psychiatrisch erfahrenste ist (15 Jahre), hält zwischen den beiden Arzttypen gewissermaßen eine Mittelstellung. Sein Urteil stabilisiert sich relativ schnell, die Entscheidung beim Anstreichen der Skalen fällt ihm jedoch wie Arzt 1 und 5 in Urteil 1 und 2 schwer.

Da insgesamt bei 3 von 5 Ärzten eine schnelle Stabilisierung des Urteils erfolgt, war das unter "Allgemeine Gesichtspunkte" referierte Ergebnis zustande gekommen. Die Unterschiede zwischen den beiden Arzttypen können unter Berücksichtigung der Ergebnisse unter "Allgemeine Gesichtspunkte" näher bestimmt werden. So betrifft die Urteilsänderung vorrangig nur die Patiententypen, die Skalenfaktoren dürften bei allen Ärzten relativ konstant bleiben.

Ein gewisses Maß an Unsicherheit bei der psychologischen Beurteilung von Patienten, insbesondere bei wenig Information, ist einem sich fortlaufend erneuernden psychologischen Bild über Patienten förderlich. Dagegen hat die Einstellung, die psychologische Beurteilung von Patienten bereite wenig Schwierigkeiten, ein sich schnell erhärtendes Urteil über die Patienten zur Folge.

Von besonderem Interesse wäre es nun festzustellen, welches der durchschnittlichen psychologischen Urteile der Ärzte das validere ist. Dazu fehlt jedoch das Außenkriterium.

2.2 Psychiatrisches Urteil

221. Stabilität

2211. Intraindividuelle Unterschiede: Psychiater

Nur bei Arzt 1 und 2 sind die Unterschiede zwischen der Höhe der Stabilitäten 1/2, 2/3 und 1/3 signifikant (Cochran Q Test; $Q_1 = 6,61$; $p_1 < 0,05$; $Q_2 = 8,66$; $p_2 < 0,02$; $Q_3 = 2,0$; $p_3 > 0,05$; $Q_4 = 0,75$; $p_4 > 0,05$; $Q_5 = 0,28$; $p_5 > 0,05$ bei jeweils $df = 2$).

Die Unterschiede werden durch die Stabilität 1/3 hervorgerufen, die erheblich niedriger ist als die beiden anderen Stabilitäten. Damit ist zumindest bei zwei Ärzten nachgewiesen, daß die Stabilität psychiatrischer Urteile gering ist, wenn für das zweite

Urteil verhältnismäßig viel und relevante Information zur Verfügung steht.

Bei keinem der Ärzte traten jedoch - wie schon berichtet - signifikante Unterschiede zwischen der Höhe der Stabilitäten 1/2 und 2/3 auf, wodurch die Hypothese, das psychologische Urteil stabilisiere sich schneller als das psychiatrische, bestätigt werden konnte. (Binomial Test; $N_1 = 13$; $X_1 = 6$; $p > 0,05$; $N_2 = 9$; $X_2 = 2$; $p > 0,05$; $N_3 = 7$; $X_3 = 2$; $p > 0,05$; $N_4 = 7$; $X_4 = 3$; $p > 0,05$; $N_5 = 5$; $X_5 = 2$; $p > 0,05$).

Auf weitere Darstellungen und Berechnungen wird wegen des geringen N verzichtet.

2212. Interindividuelle Unterschiede: Psychiater

Die Differenzen zwischen den Ärzten bezüglich der Höhe der Stabilitäten sind weder für die Stabilitäten 1/2 und 2/3 noch für die Stabilität 1/3 signifikant (Cochran Q Test; $Q_{1/2} = 1,71$, $p_{1/2} > 0,05$; $Q_{2/3} = 2,94$, $p_{2/3} > 0,05$; $Q_{1/3} = 9,29$, $p_{1/3} > 0,05$ bei jeweils $df = 4$), doch liegen die Differenzen bei der Stabilität 1/3 nahe dem 5% Signifikanz-Niveau. Danach hat ein relativ großes Ausmaß an unterschiedlicher Information zumindest tendenziell Einfluß auf die Stabilitätshöhe der Urteile verschiedener Ärzte.

Der Vergleich der Diagnosehäufigkeit zwischen den Ärzten gibt Aufschluß über individuelle Diagnosetendenzen (Tabelle 21). Die Häufigkeiten sind jeweils über die drei Urteile summiert.

Arzt		1	2	3	4	5
Diagnose	B	17	12	6	7	10
	N	14	9	12	15	8
	S	7	13	21	9	16
	Z	6	9	3	11	6

Tabelle 21: Gesamtzahl der Beschreibungen und Diagnosen N, S und Z pro Arzt

Eine statistische Überprüfung auf interindividuelle Unterschiede soll wegen der Summation über die drei Urteile unterlassen werden. Für eine entsprechende Berechnung pro Urteil sind die Häufigkeiten zu gering. Aus der Tabelle geht jedoch deutlich hervor, daß erhebliche interindividuelle Unterschiede in der Anzahl der B und der aller drei

Diagnosegruppen bestehen. So haben beispielshalber Arzt 3 eine Vorliebe für S-, Arzt 1 und 4 eine Vorliebe für N-Diagnosen. Arzt 1 scheint der unsicherste der 5 Ärzte zu sein, da er die insgesamt größte Anzahl von B abgab.

2213. Differentielle Aspekte: Patienten

Die in diesem Fall geringe Anzahl von Merkmalsträgern (5 Ärzte) gestattet keine der üblichen statistischen Überprüfungen auf intra- oder interindividuelle Differenzen bezüglich der Stabilität der Beurteilbarkeit von Patienten.

Die Größe der Stabilität der Beurteilbarkeit zwischen den Urteilen kann jedoch folgendermaßen überprüft werden, wobei jedoch eine Vermischung der Merkmalsträger Ärzte und Patienten in Kauf genommen wird: Die Anzahl der von den 5 Ärzten pro Patient abgegebenen stabilen Urteile von Urteil 1 nach 2, 2 nach 3 und 1 nach 3 wird ermittelt. Diese Häufigkeitswerte gelten als Maß für die jeweilige Beurteilbarkeitsstabilität pro Patient.

Mit Hilfe des Friedman Tests werden diese 3 Stabilitäten auf Unterschiede bezüglich ihrer zentralen Tendenz überprüft. Die Differenzen sind nicht signifikant ($\chi^2 = 3,42$; $df = 2$; $p > 0,05$).

Für die Stabilitäten 1/2, 2/3 und 1/3 sind mit der oben genannten Vorgehensweise die Variablen Stabilität 1/2', 2/3', 1/3' gewonnen. Die Ranginterkorrelation dieser Variablen über die Patienten gibt nun Aufschluß über die relative Stabilität der Beurteilbarkeit von Patienten (Tabelle 22).

Bei der Berechnung der Rangkorrelationen werden gleiche Ränge der Meßwerte durch Korrektur berücksichtigt (tied observation).

Variable Stabilität	2/3'	1/3'
1/2'	-.234	-.054
2/3'		.372

Tabelle 22: Beziehung zwischen den Variablen Stabilität 1/2', 2/3' und 1/3' (Rangkorrelationskoeffizienten) $r_s \geq .43$, $p \leq 0,05$; $r_s \geq .63$, $p \leq 0,01$

Die mit Hilfe des Cochran Q Tests überprüften Unterschiede zwischen den drei Urteilen bezüglich der Anzahl von R für alle Patienten sind bei den Ärzten 1 und 4 hoch signifikant, bei den Ärzten 2, 3 und 5 nicht signifikant. Entsprechende Berechnungen bei den B ergeben für Arzt 1 eine hoch signifikante Differenz, für die Ärzte 2, 4 und 5 signifikante Differenzen und für Arzt 3 keine signifikante Differenz. Bezüglich der Anzahl von F tritt bei keinem der Ärzte eine signifikante Differenz auf (Tabelle 23).

Arzt	R	B	F
1	9,75	12,18	2,00
2	3,20	7,00	0,25
3	4,00	3,60	1,20
4	10,89	7,60	3,60
5	2,00	7,00	3,00

Tabelle 23: Q-Werte; Erläuterung im Text. df jeweils = 2; $Q \geq 5,99$, $p \leq 0,05$; $Q \geq 9,21$, $p \leq 0,01$

Anwachsende und psychiatrisch relevanter werdende Information führt bei 4 von 5 Ärzten zur statistisch abgesicherten Abnahme der Anzahl von B, aber nur bei 2 der 5 Ärzte zur statistisch abgesicherten Zunahme der Anzahl von R. Informationsausmaß und - art nimmt bei keinem der Ärzte signifikanten Einfluß auf die Anzahl von F.

Das Ausmaß der subjektiven Diagnosesicherheit - eingeschätzt anhand der Skala Diagnose: sicher vs. sehr unsicher - unterscheidet sich bei allen 5 Ärzten zwischen den 3 Urteilen signifikant, und zwar steigt die Diagnosesicherheit beständig an (Friedman Test; $\chi^2_1 = 18,63$, $p_1 < 0,001$; $\chi^2_2 = 14,93$, $p_2 < 0,001$; $\chi^2_3 = 12,90$, $p_3 < 0,01$; $\chi^2_4 = 9,70$, $p_4 < 0,01$; $\chi^2_5 = 24,03$, $p_5 < 0,001$ bei jeweils df = 2).

Da nur bei 2 der 5 Ärzte eine signifikante Zunahme der Anzahl von R, jedoch aber bei 4 der 5 Ärzte ein signifikanter Abfall der Anzahl von B auftritt, scheint die selbstbeurteilte Diagnosesicherheit eher Ausdruck für das emotional entlastende Gefühl zu sein, überhaupt eine Diagnose anbieten zu können, als für die Diagnosesicherheit selbst (vgl. RANABAUER, 1968).

Dieses Ergebnis ist jedoch insofern banal, als bei Beschreibungen völlige Diagnoseunsicherheit angestrichen wurde, während sich bei Diagnosen die eingestufte Diagnosesicherheit erhöhte.

Weitere statistische Differenzüberprüfungen, wie Unterschiede der Anzahlen pro R, F und B zwischen den Urteilen bei jeder der 3 Diagnosegruppen oder Unterschiede der Anzahlen pro R, F und B zwischen den Diagnosegruppen bei jedem der 3 Urteile, können wegen des dann zu geringen N von 5 nicht vorgenommen werden. Auch eine graphische Darstellung erscheint nicht sinnvoll.

2222. Interindividuelle Unterschiede: Psychiater

Bei der interindividuellen statistischen Differenzüberprüfung der Anzahlen pro R, F und B und pro Urteil müssen alle Patienten als Meßwertträger berücksichtigt werden; eine entsprechende Überprüfung pro Diagnosegruppe ist auch hier wegen des geringen N von nur 5 nicht mehr zu rechtfertigen.

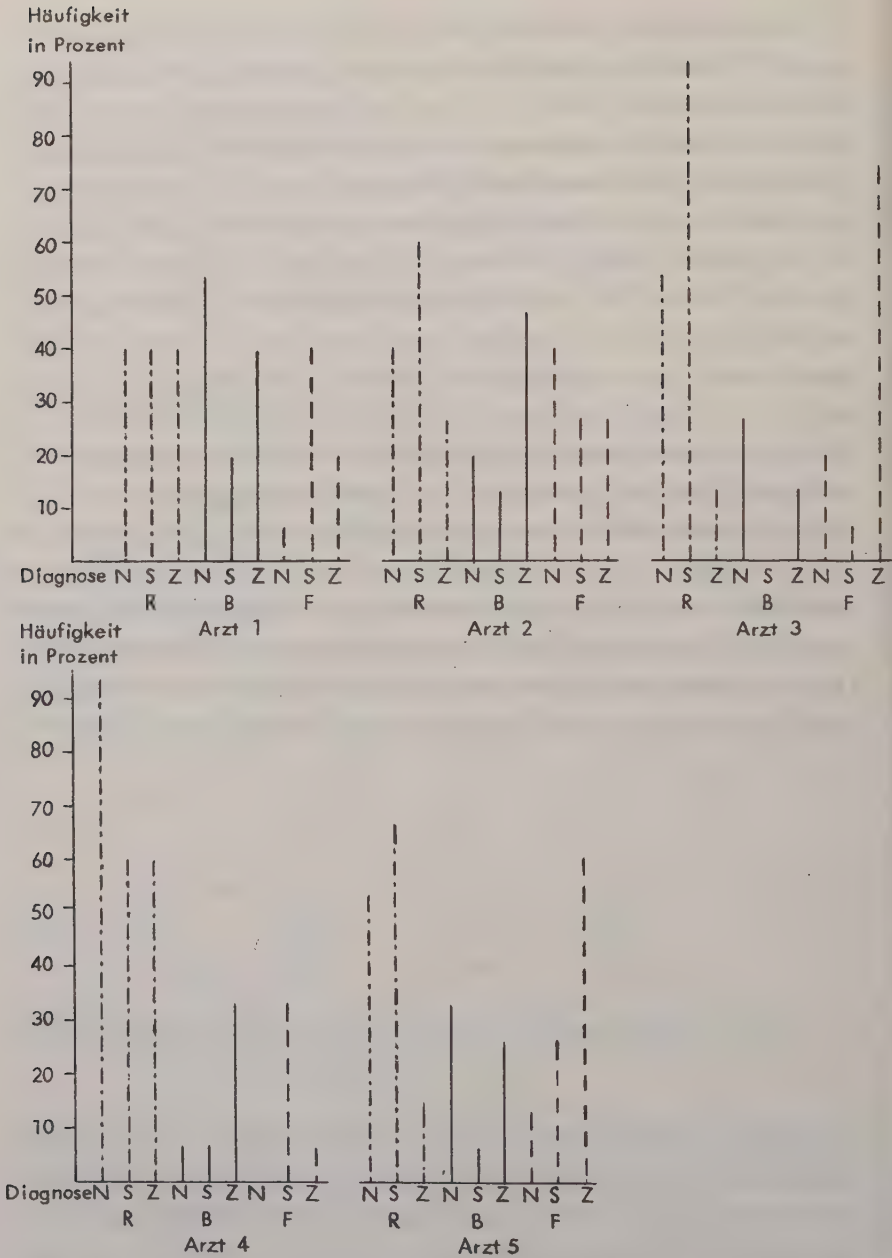
Während sich die Ärzte bei Urteil 2 weder bezüglich der Anzahlen von R und F noch bezüglich der von B signifikant unterscheiden, treten bei Urteil 1 bezüglich der Anzahlen von B und bei Urteil 3 jeweils bezüglich der Anzahlen von R und F signifikante Unterschiede auf (Tabelle 24).

Urteil	1	2	3
	Q	Q	Q
R	4,67	4,12	10,62
B	9,67	2,26	5,33
F	2,93	0,69	9,93

Tabelle 24: Q-Werte; Erläuterung im Text. df jeweils = 4; $Q \geq 9,49$; $p \leq 0,05$; $Q \geq 13,28$; $p \leq 0,01$

Im Ausmaß geringe und psychiatrisch nur minimal relevante Information bedingt somit signifikante Unterschiede zwischen den Ärzten bezüglich des Grades der Urteilssicherheit, im Sinne des diagnostischen Festlegens vs. Beschreibens. In diesem Stadium des Diagnoseprozesses treten noch keine signifikanten Unterschiede bezüglich der psychia-

trisch-diagnostischen Fähigkeiten, im Sinne des richtigen bzw. falschen Diagnostizierens auf. Diese Unterschiede machen sich erst nach viel und psychiatrisch wichtiger Information signifikant bemerkbar. Der valide diagnostizierende Psychiater unterscheidet sich damit vom weniger valide diagnostizierenden nicht durch ein besonderes diagnostisches Gespür, das ihm nach schon kurzem Kontakt mit dem Patienten die richtige Diagnose zu fällen erlaubt, sondern durch korrekte Verarbeitung von viel und psychiatrisch wichtiger Information. Die Vermutung, daß zwischen Urteilsunsicherheit im ersten Urteil und diagnostischer Fähigkeit bzw. Unfähigkeit im dritten Urteil ein negativer bzw. positiver Zusammenhang besteht, konnte nur tendenziell bestätigt werden. So betragen die über die Ärzte berechneten Rangkorrelationen zwischen den B aus dem 1. Urteil und den R bzw. F aus dem 3. Urteil $r = -.657$ bzw. $r = .026$. Beide Koeffizienten sind nicht signifikant. Als Meßwerte gingen die pro Arzt über alle Patienten summierten B, R und F in die Berechnung ein. Weitere statistische Berechnungen sind wegen des dann jeweils zu geringen N von nur 5 nicht möglich. Graphisch dargestellt werden sollen jedoch die interindividuellen Unterschiede in Bezug auf Urteilsunsicherheit (B) und Urteilsfähigkeit (R) bzw. -unfähigkeit (F). Die veranschaulichten prozentualen Anteile basieren auf den über die 3 Urteile summierten Häufigkeiten. Die graphische Darstellung 9 kann auch nach intraindividuellen Gesichtspunkten interpretiert werden.



Graph. Darst. 9: Prozentuale Anteile aller richtigen und falschen Diagnosen und Beschreibungen pro Diagnosegruppe für jeden Arzt

Auffällig sind insbesondere der hohe Anteil richtiger S-Diagnosen bei Arzt 3 und richtiger N-Diagnosen bei Arzt 4. Dieser Sachverhalt kann jedoch nicht eindeutig als ein krankheitsgruppenspezifisches Gespür bzw. als eine differentielle Urteilsfähigkeit interpretiert werden. Denn bei Arzt 3 ist die absolute Anzahl von S-Diagnosen und ebenfalls S-Falschdiagnosen am höchsten. Und Arzt 4 zeigt ein entsprechendes Übergewicht an N-Diagnosen und Falschdiagnosen in Richtung Neurose, Entwicklungskrise. Daher erscheint es eher sinnvoll, von einer Vorliebe für bestimmte Diagnosen zu sprechen, womit die Interaktion zwischen den beiden Merkmals-trägern beurteilte Patienten und beurteilende Ärzte angedeutet ist.

Während die hier aufgezeigten individuellen Diagnosetendenzen - ähnlich wie unter allgemeinen Gesichtspunkten - durch Bezugnahme auf die pro Diagnosegruppe insgesamt abgegebenen Diagnosen z.T. relativiert werden können - was hier wegen der geringen Häufigkeiten nicht geschehen soll - lassen sich diese Diagnosetendenzen - im Sinne negativer Diagnosetendenzen - jedoch eindeutig anhand der Falschdiagnosen nachweisen. So diagnostiziert Arzt 3 Patienten der Gruppe Z am häufigsten falsch, gibt aber nur eine nicht zutreffende Z-Diagnose.

Arzt 1 zeigt eine Diagnosetendenz in Richtung N-Diagnosen, der Anteil richtiger N-Diagnosen ist jedoch gegenüber dem richtiger S- und Z-Diagnosen nicht erhöht, und entsprechend haben N-Falschdiagnosen den höchsten Anteil aller Falschdiagnosen. Dieser Arzt hat offensichtlich trotz Vorliebe für N-Diagnosen ein besonders schlechtes diagnostisches Gespür für Patienten dieser Diagnosegruppe.

Weitere, bisher nur exemplarisch dargestellte Zusammenhänge können aus dem Vergleich zwischen der graphischen Darstellung 9 und folgender Tabelle 25 ermittelt werden.

Arzt	1	2	3	4	5
	80% N	36% Z	47% S	33% N	40% S
	10% S	29% S	26% N	33% Entw. Krise	27% Z
	10% endo- reakt.	21% N	13% Psy- chopath	33% Z	20% Psychopath
	Depre.	14% Psy- chopath	7% Z		13% endo-reakt. Depression
			7% endo- reakt. Depres.		

Tabelle 25: Prozentualer Anteil aller falschen Diagnosen pro Arzt

Beim psychologischen Urteil waren 2 Arzttypen herausgestellt worden. Die Ärzte 1 und 5 fühlten sich beim Urteilen anfänglich erheblich unsicher, und die stärkste Urteilsänderung trat bei ihnen von Urteil 2 nach 3 auf. Im Gegensatz dazu bereitet den Ärzten 2 und 4 die psychologische Beurteilung auch anfangs wenig Schwierigkeiten, und ihre Beurteilung stabilisierte sich nach dem 2. Urteil. Interpretativ wurde zusammengefaßt, daß die Ärzte 1 und 5 im Verhältnis zu den Ärzten 2 und 4 den Schwerpunkt ihres Urteils auf eine psychologische Beurteilung der Patienten legten.

Beim psychiatrischen Urteil findet sich ein in etwa entsprechender Unterschied zwischen den beiden Arzttypen wieder. Zwar erweist sich der Abfall des Ausmaßes subjektiver Diagnosesicherheit bei allen 4 Ärzten als signifikant, jedoch ist der graphischen Darstellung 8 zu entnehmen, daß die Ärzte 1 und 5 die psychiatrisch besonders relevante Information aus Situation IV und V weniger effektiv verarbeiteten als die Ärzte 2 und 4. Der Anstieg des Anteils der R ist bei den Ärzten 1 und 5 von Urteil 1 nach 2 am stärksten, bei den Ärzten 2 und 4 von Urteil 2 nach 3. Arzt 3 beweist auch hier seine Mittelstellung; der Anstieg des Anteils der R von Urteil 1 bis 3 ist linear.

2223. Differentielle Aspekte: Patienten

Die üblichen statistischen Differenzüberprüfungsverfahren zur Ermittlung von intra- und interindividuellen Unterschieden können wegen der geringen Anzahl von Meßwertträgern (5 Ärzte) nicht angewendet werden.

Folgendermaßen können jedoch - in Analogie zum Vorgehen unter Stabilität des Urteils Unterschiede im Ausmaß der Beurteilbarkeit in Urteil 1, 2 und 3 überprüft werden. Die Anzahl der von den 5 Ärzten zu jedem Patienten getroffenen R, F und B wird pro Urteil ermittelt. Die Häufigkeitswerte der R bzw. F eines jeden Patienten gelten als Maß für die Beurteilungsleichtigkeit bzw.-schwierigkeit des Patienten, die Häufigkeitswerte B als Maß für die Beurteilungsmehrdeutigkeit. Überprüft wird nun mit Hilfe des Friedman Tests, ob pro R, B und F Unterschiede in der zentralen Tendenz zwischen den Urteilen bestehen. Die Differenzen sind bei R signifikant ($\chi^2 = 6,63$; bei $df = 2$; $p < 0,05$), bei den B sehr hoch signifikant ($\chi^2 = 15,73$ bei $df = 2$; $p < 0,001$), bei den F treten keine signifikanten Differenzen auf ($\chi^2 = 0,32$ bei $df = 2$; $p > 0,05$), d.h. die Beurteilungsleichtigkeit nimmt von Urteil 1 bis 3 signifikant zu, die Beurteilungsmehr-

deutigkeit hoch signifikant ab.

Pro Analyse sind über das obige Vorgehen die Variablen R' , F' und B' gewonnen. Diese Variablen werden jeweils zwischen den Urteilen ranginterkorreliert, wobei durch Korrektur Meßwerte mit demselben Rang berücksichtigt werden (Tabelle 26).

Urteil	1/2	1/3	2/3
Variable			
R'	.483	-.113	.196
B'	.340	.016	.525
F'	.195	.546	.113

Tabelle 26: Zusammenhang zwischen den 3 Urteilen für Beurteilungsleichtigkeit, -mehrdeutigkeit und -schwierigkeit (Rangkorrelationskoeffizienten)
 $r_s \geq .43$, $p \leq 0,05$; $r_s \geq .63$, $p \leq 0,01$

Die Beurteilungsleichtigkeit der Patienten bleibt vom 1. zum 2. Urteil signifikant konstant, zum 3. Urteil verschwindet sie. Umgekehrt ist der Sachverhalt bei der Beurteilungsmehrdeutigkeit. Zusammenfassend können die Ergebnisse dahingehend interpretiert werden, daß die Information aus Situation II und III entscheidenden Einfluß auf die Beurteilungsmehrdeutigkeit, die zusätzliche Information aus Situation IV und V auf die Beurteilungsleichtigkeit nimmt. Die Information aus Situation II und III ändert somit vorrangig die diagnostische Mehrdeutigkeit der Patienten, während die zusätzliche Information aus Situation IV und V zu einem Wechsel der Beurteilungsleichtigkeit beiträgt.

Interessant ist die Konstanz der Beurteilungsschwierigkeit von 1. zu 3. Urteil. Das Verhältnis der Beurteilungsschwierigkeit zwischen den Patienten verschiebt sich aufgrund der Information aus Situation II und III, aufgrund der weiteren Informationen aus Situation IV und V wird aber das ursprüngliche Verhältnis wiederhergestellt. Aufgrund des ersten Eindrucks schwer bzw. leicht diagnostizierbare Patienten bereiten im letzten Urteil entsprechend große bzw. geringe Schwierigkeiten. Dieses Ergebnis legt die Vermutung nahe, daß es schwer und leicht diagnostizierbare Patienten gibt. Relativ wenig und psychiatrisch nicht sehr ergiebige Information scheint den Grad ihrer Beurteilungsschwierigkeit zu verschleiern.

Unterschiede zwischen den 3 Diagnosegruppen bezüglich des Ausmaßes der 3 Komponenten der Beurteilbarkeit können mit Hilfe des Kruskal Wallis Test statistisch überprüft werden (Korrektur bei tied observation). Die 3 Diagnosegruppen unterscheiden sich nur im ersten Urteil, bezüglich des Ausmaßes ihrer Beurteilungsleichtigkeit und -schwierigkeit signifikant, also nur bei sehr wenig und psychiatrisch irrelevanter Information (Tabelle 27).

Urteil	1	2	3
R	8,75	3,61	2,04
B	4,18	6,91	2,33
F	9,25	1,96	1,66

Tabelle 27: H-Werte; Erläuterung im Text. jeweils $df = 2$; $H \geq 5,78$, $p \leq 0,049$
 $H \geq 7,98$, $p \leq 0,010$

Durch Zunahme von relevanter Information werden diese Unterschiede aufgehoben.

Die Information aus Situation II und III führt im 2. Urteil zu signifikanten Unterschieden in der Beurteilungsmehrdeutigkeit.

In Urteil 3 findet ein Ausgleich in der Beurteilbarkeit statt; es treten keine signifikanten Differenzen auf.

Die Rangkorrelationen pro Urteil zwischen den Patientenkommunalitäten und jeweils den Variablen R', F' und B' geben Aufschluß über den Zusammenhang zwischen Bestimmtheit im psychologischen Urteil und Beurteilbarkeit bzw. Bestimmbarkeit im psychiatrischen Urteil. Keine der 9 Rangkorrelationen ist auch nur annähernd signifikant, d.h. es gibt keinen allgemeinen Zusammenhang zwischen psychologischer und psychiatrischer Beurteilbarkeit der Patienten.

Der per inspectionem vorgenommene Vergleich der pro Diagnosegruppe summierten Kommunalitäten und Beurteilungswerte zwischen den Diagnosegruppen zeigt jedoch auf, daß sich die schizophrenen Patienten von den neurotischen und cyclothymen durch größere Bestimmtheit im psychologischen und erhöhte Beurteilungseindeutigkeit (wenig B) im psychiatrischen Urteil in allen 3 Urteilen unterscheiden.

II Psychologische Beurteilung der Psychiater

Jeder Arzt explorierte insgesamt 15 mal. Da jeder Arzt Beurteiler und Beurteilter war, gab er jeweils 15 Beurteilungen über seine 4 Kollegen und sich selbst ab und erhielt jeweils von seinen 4 Kollegen und sich selbst 15 Beurteilungen. Als Beurteilungsinstrument diente das Arzt-Polaritäten-Profil.

Die Explorationsabfolge der Patienten ist durch das lateinische Quadrat als Untersuchungsplan festgelegt. Somit erfolgte die Beurteilung des Explorationsverhaltens eines jeden Arztes jeweils nach jeder der 5 Explorations-Situationen 3 mal, nämlich nach Exploration eines Neurotikers, eines Schizophrenen und eines Cyclothymen. Die 15-maligen Beurteilungen entsprechen sich aber nicht genau zwischen den Ärzten, da laut lateinischem Quadrat jeder Arzt zwar jeweils denselben Patienten, nicht aber in derselben Situation exploriert. Daher können die Beurteilungen der Ärzte nach folgenden beiden Kriterien geordnet werden.

a) Die Beurteilung des Explorationsverhaltens eines Arztes wird bestimmt durch die Diagnosegruppe des explorierten Patienten und die Explorationssituation. Die Beurteilungen des Explorationsverhaltens der Ärzte entsprechen sich dann insofern, als jeder Arzt nach Exploration von Patienten einer bestimmten Diagnosegruppe in einer bestimmten Situation eingeschätzt wird. Somit erfolgen die Beurteilungen also nicht nach Exploration desselben Patienten, sondern verteilen sich entsprechend dem lateinischen Quadrat über die Patienten einer Diagnosegruppe. So ist die 1. Beurteilung des Arztes 1 die Beurteilung dieses Arztes nach Exploration des ersten Neurotikers in Situation 1, dagegen bei Arzt 2 die Beurteilung dieses Arztes nach Exploration des dritten Neurotikers in Situation 1.

b) Die Beurteilung des Explorationsverhaltens eines Arztes wird gekennzeichnet durch einen bestimmten Patienten, der einer der 3 Diagnosegruppen angehört. Die Explorationssituation ist unbestimmt, die Explorations-Situationen verteilen sich entsprechend dem lateinischen Quadrat auf die Beurteilung der Ärzte nach Exploration der Patienten einer Diagnosegruppe.

Da jeder Arzt Beurteiler und Beurteilter ist, kann er sowohl als Beurteiler als auch Beurteilter analysiert werden.

Die Analyse des Arztes als Beurteiler kann in zweierlei Hinsicht erfolgen:

1) Die jeweils 15 Beurteilungen, die er über seine 4 Kollegen und sich selbst trifft, werden jeweils entsprechend a) geordnet, d.h. jeweils jede der 15 Beurteilungen ist bestimmt durch eine Diagnosegruppe und eine Explorationssituation.

2) Die jeweils 15 Beurteilungen werden jeweils entsprechend b) gekennzeichnet, d.h. jeweils jede der 15 Beurteilungen ist durch einen der 15 Patienten bestimmt.

Bei der Analyse des Arztes als Beurteiltem kann ebenfalls im Sinne 1) und 2) eine Unterscheidung vorgenommen werden.

Als Analysemethode bietet sich in allen Fällen wiederum die 3 Modi-FA an. Die Modi sind unter der Bedingung Arzt als Beurteiler a) Skalen, b) Beurteilte, c) 15-malige Beurteilungen und unter der Bedingung Arzt als Beurteilter a) Skalen, b) Beurteiler, c) 15-malige Beurteilung.

Das Explorationsverhalten eines jeden Arztes wird grob durch folgende Faktoren beeinflusst:

- a) die gruppenspezifischen Verhaltensweisen von Patienten einer bestimmten Diagnosegruppe
- b) die Explorations-Situation und
- c) die jeweils spezifischen Verhaltensweisen der individuellen Patienten.

Bei Faktor a) gilt als Determinante des Explorationsverhaltens die gruppenspezifische Verhaltenseigentümlichkeit der Patienten einer bestimmten Diagnosegruppe; das spezifische Verhalten des individuellen Patienten bleibt dabei unberücksichtigt.

Unter Faktor b) bedingt das Explorationsthema pro Situation den Explorationsstil und unter Faktor c) beeinflusst das spezifische Verhalten des individuellen Patienten die Exploration insofern, als sich der jeweils explorierende Psychiater auf den jeweiligen Patienten einstellen muß.

Der Einfluß der Faktoren a) und b) auf das Explorationsverhalten ist durch die Analyseform 1 (Bestimmung der Beurteilungen durch Diagnosegruppe und Explorations-Situation) und der Effekt des Faktors c) durch Analyseform 2 (Bestimmung der Beurteilungen durch den individuellen Patienten) abschätzbar. Der Effekt dieser Faktoren wird jedoch nur über das Urteil der Psychiater anhand der Beurteilungsskalen ermittelt, da kein exaktes Außenkriterium vorliegt. Eindeutige Aussagen über die Wirkung der Faktoren auf das reale Explorationsverhalten können daher nicht getroffen werden. Der Einfluß der Faktoren wird über die im Urteil implizierte Ansicht der Psychiater ermittelt, wobei individuelle Urteilstendenzen, aber auch mögliche Störvariablen wie die Überforderung, bei schwierigen Explorationen auch noch den explorierenden Arzt zu beobachten und zu beurteilen, von vornherein zu einer Verzerrung der realen Sachverhalte führt.

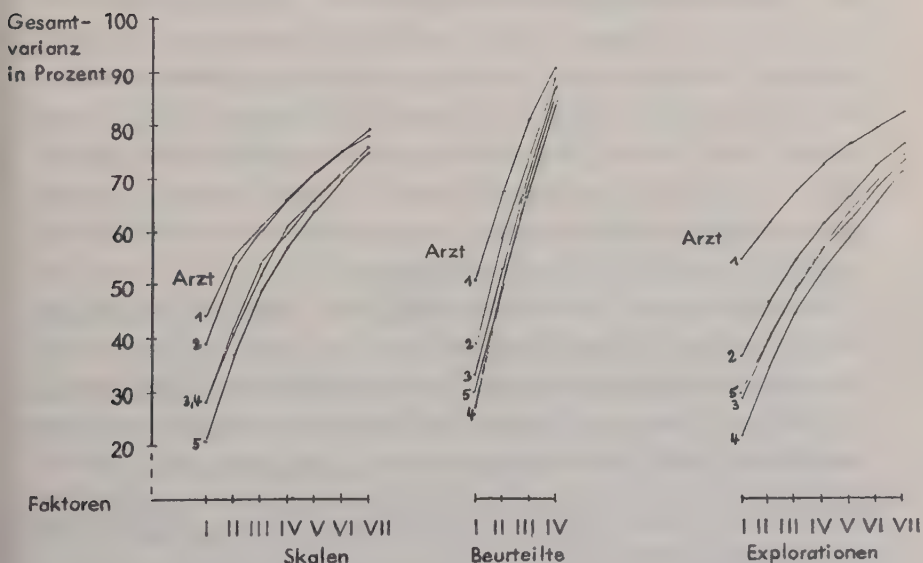
1. Ärzte als Beurteiler

11. Bestimmungen der Urteile durch Diagnosegruppe und Explorationssituation.

Pro beurteilendem Arzt wird eine 3 Modi-FA durchgeführt. Die Modi sind Skalen, beurteilte Ärzte (4 Kollegen, der urteilende Arzt selbst) und die 15-fachen Urteile. In den letzten Modus wird als 16. Variable die aus der Vorstellung heraus gegebene Beurteilung des typischen Verhaltens der Ärzte aufgenommen.

111. Differenziertheit der Urteile

Die Varianzanteile der unrotierten Faktoren pro Modus sind für jede der 5 Arzt-Analysen in folgender graphischer Darstellung 10 wiedergegeben.



Graph. Darst. 10: Kumulierte prozentuale Gesamtvarianzanteile der unrotierten Faktoren pro Modus und beurteilendem Arzt

Beim Skalen-Modus liegen insbesondere in der Varianzstärke des jeweils 1. Faktors erhebliche interindividuelle Unterschiede vor. Auch die Anzahl der als bedeutsam anzusehenden Faktoren variiert. Bei Arzt 3, 4 und 5 erscheinen jeweils 4 Faktoren als bedeutsam, bei Arzt 1 und 2 jeweils nur 2 Faktoren. Demnach differiert der Differenz-

ziertheitsgrad der individuellen kognitiven Beurteilungsräume bei der Arztbeurteilung ganz beträchtlich. Die Unterschiede sind sowohl bezüglich der Stärke des jeweils 1. Faktors als auch bezüglich der Anzahl der jeweils bedeutsamen Faktoren weitaus größer als die entsprechenden interindividuellen Differenzen bei der Patientenbeurteilung. Diese Diskrepanz kann grob durch 3 Faktoren hervorgerufen werden:

- a) Die Psychiater werden in ihrer Ausbildung auch im psychologischen Beurteilen von Patienten geschult. Folge dieser Schulung ist ein in etwa gemeinsames Wissen und psychiatrisches Können, das zu einer Reduzierung interindividueller Differenzen im Differenziertheitsgrad der Urteilsstrukturen führt. Eine weitere Verminderung der Unterschiede resultiert aus der gemeinsamen Arbeit an Patienten; durch Diskussion wird - auch bezüglich der Differenziertheit des Urteils - über den Zwang zur Konformität Konsensus erreicht. Im Gegensatz dazu ist die Beurteilung des Explorationsverhaltens von Psychiatern ungeübt, so daß interindividuelle Differenzen voll zur Wirkung kommen.
- b) Als wesentliche Aufgabe galt für die Psychiater in der Untersuchung die Beurteilung von Patienten. Die Leistungsmotivation in Bezug auf diese Aufgabe dürfte bei allen Ärzten gleich gewesen sein. Die zusätzliche Aufgabe, nämlich Beobachtung und Beurteilung der Ärzte, wurde als relativ unwichtig betrachtet und überforderte möglicherweise in interindividuell unterschiedlichem Ausmaß die Ärzte. Unterschiede im Grad der Urteilsdifferenziertheit sind daher auf interindividuelle Differenzen im Ausmaß der Leistungsmotiviertheit und Überforderung zurückzuführen.
- c) Die Beurteilungsskalen sind zur Einschätzung des Explorationsverhaltens ungeeignet. Sie decken unwesentliche Komponenten des Explorationsverhaltens ab. Bezüglich der differenzierten Verwendungsfähigkeit dieser Skalen bestehen interindividuelle Differenzen. Diese Unterschiede sagen jedoch nur noch wenig über Differenzen im Ausmaß der Urteilsdifferenziertheit bei der Einschätzung von Explorationsverhalten aus.

Faktor b) und c) können sicherlich nicht unberücksichtigt bleiben, Faktor a) erscheint jedoch als der plausibelste.

Auch beim Beurteilten-Modus liegen große interindividuelle Unterschiede bezüglich der Varianzstärke des ersten Faktors sowie bezüglich der Anzahl bedeutsamer Faktoren vor. So gibt es z.B. im Urteil von Arzt 1 nur einen Beurteiltentyp, also Arzttyp, während Arzt 4 die 5 Ärzte deutlich zu 2 Typen ordnet. Bei den restlichen 3 Ärzten scheinen ebenfalls 2 Faktoren bedeutsam zu sein.

Entsprechende interindividuelle Unterschiede treten auch beim Beurteilungsmodus auf. Die Ärzte unterschieden sich beträchtlich in ihrer Fähigkeit, die 16 Explorationen in ihrem Urteil zu differenzieren. So faßt Arzt 1 die 16 Explorationen zu einem Faktor zusammen, die Ärzte 2, 3 und 4 zu 3 Faktoren, während Arzt 5 sogar annähernd eine Ordnung in 4 Faktoren vornimmt.

Die oben genannten 3 Faktoren können ebenfalls zur Erläuterung der starken interindividuellen Differenzen beim Beurteilten- und Beurteilungs-Modus herangezogen werden.

112. Skalen

Trotz der großen interindividuellen Unterschiede in der Anzahl der als bedeutsam anzusehenden Skalenfaktoren wurden nach genauer Analyse der Eigenwerte jeweils 4 Faktoren extrahiert und zur Interpretation varimaxrotiert. Außerdem wurden wie üblich die 5 Faktorenmatrizen auf ihre Ähnlichkeit hin verglichen.

Die Ähnlichkeit der Faktorenstrukturen ist minimal. Das Mittel der Koeffizienten beträgt .593. Die Varianz der Koeffizienten ist sehr gering (.094), entsprechend erreicht die größte Ähnlichkeit von Faktorenstrukturen nur einen Wert von .716. Zwar werden die geringen Faktorenstrukturähnlichkeiten sicherlich auch durch die z.T. zu große Anzahl von 4 Faktoren, die in die Vergleiche eingingen, mitbestimmt, doch scheint dieser Sachverhalt nicht von ausschlaggebendem Einfluß zu sein. Denn die Faktorenstrukturähnlichkeit zweier Matrizen erhöht sich nicht, wenn jeweils alle 4 Faktoren signifikant sind. So erscheint die Interpretation gerechtfertigt, daß die kognitiven Beurteilungsstrukturen bei der Beurteilung von explorierenden Ärzten nur noch sehr geringe inhaltliche interindividuelle Übereinstimmungen aufweisen. Diese Unterschiede sind bei weitem größer als die entsprechenden bei der Patientenbeurteilung. Zur Erläuterung dieser Diskrepanz können die schon unter dem Aspekt der Differenziertheit des Urteils genannten Faktoren herangezogen werden.

Es bestehen beträchtliche Unterschiede im Grad der Interpretationsschwierigkeit der 5 Skalenfaktorenmatrizen. Dies hängt z.T. damit zusammen, daß bei einigen Ärzten, insbesondere bei Arzt 1, die Anzahl der extrahierten und rotierten Faktoren größer ist als die signifikanten Faktoren. Im folgenden soll lediglich eine, nämlich die Faktorenmatrix des Arztes 4, interpretiert werden, um eine der aufgetretenen Urteilsstrukturen zu veranschaulichen. Die Interpretation der Faktoren dieser Matrix fällt relativ

leicht; bei Arzt 4 können 4 Faktoren als signifikant gelten. Wie üblich, sind Skalen mit einer Ladung gleich/größer .250 aufgeführt und Skalen mit einem negativen Vorzeichen invertiert, die negativen Vorzeichen jedoch belassen (Tabelle 28).

Faktor

I	18)	.440	beharrend	-	schnell aufgebend
	11)	-.417	mißtrauisch	-	arglos
	10)	-.375	lebhaft	-	still
	15)	-.358	schnell	-	langsam
	1)	.289	aktiv	-	passiv
	5)	.250	im Verhalten ausdrucksvoll	-	im Verhalten farblos
II	14)	-.578	sachlich	-	selbstbezogen
	2)	.498	rational	-	gefühlsmäßig
	7)	-.320	selbstkritisch	-	selbstgefällig
	21)	-.307	gleichbleibend	-	wechselnd
	6)	-.255	zielgerichtet	-	weitschweifig
III	11)	.434	arglos	-	mißtrauisch
	12)	-.417	gewährend	-	streng
	9)	-.363	gelassen	-	irritierbar
	20)	-.344	gelöst	-	gespannt
	16)	-.338	unbefangen	-	kontrolliert
IV	13)	.723	originell vorgehend	-	schematisch vorgehend
	6)	.351	weitschweifig	-	zielgerichtet
	4)	.295	geduldig	-	ungeduldig
	19)	.267	interessiert	-	uninteressiert

Tabelle 28: 4 rotierte Skalenfaktoren des urteilenden Arztes 4

In fast allen Skalen werden sowohl persönlichkeitspezifische als auch explorationsstilistische Verhaltenskomponenten angesprochen, jedoch jeweils mit unterschiedlichem Gewicht.

In jedem Faktor treten beide Skalenarten auf. Da die Faktoren jeweils als Einheit interpretiert werden, wird die Trennung zwischen den beiden Skalen bei der Interpretation nicht durchgängig beibehalten. Die Bipolarität bleibt in der Interpretation unberücksichtigt, um die Faktoren verständlicher zu machen.

Faktor I enthält die Komponenten "Ausdauer" und "Dynamik" im explorativen Verhalten. Die Dimension "Problemzentriertheit", d.h. "Patientenzentriertheit" kennzeichnet den zweiten Faktor. Faktor III ist charakterisiert durch die persönlichkeits-spezifische Komponente "Ausgeglichenheit", mit der eine "gewährende Haltung" gegenüber Patienten als eher explorationsstilistischer Verhaltensaspekt einhergeht. In Faktor IV kommen die Komponente "Explorationstechnik" und die mit ihr gekoppelte Komponente "Einstellung zum Patienten" zum Ausdruck. Es kann als sicher gelten, daß allen Faktoren Bewertungsaspekte anhaften, denn die ungeübte Beurteilung des psychiatrischen Explorationsverhaltens erfolgt mit großer Wahrscheinlichkeit nach dem Gesichtspunkt "gut vs. schlecht". Entsprechend bemängelten auch die an der Untersuchung teilnehmenden Ärzte das Fehlen von Skalen, die eine eindeutig bewertende Stellungnahme bezüglich der Güte des Explorationsverhaltens gestatten.

113. Beurteilte

Es wurden 2 Faktoren extrahiert und die üblichen Rechenschritte vorgenommen. Der Mittelwert der Ähnlichkeitskoeffizienten beträgt .652, die Streuung .126. Die relativ große Streuung weist darauf hin, daß sich die Beurteilungstypen zwischen den Ärzten z.T. einigermaßen gut, z.T. gar nicht ähneln. So weisen z.B. die Beurteiltentypen von Arzt 2 und 5 eine deutliche Ähnlichkeit auf (.864), bei den Beurteiltentypen von Arzt 1 und 4 fehlt fast jede Ähnlichkeit (.470).

Da kein direkter Vergleich zwischen Fremd- und Selbstbeurteilung vorgenommen wird, soll das Verhältnis der Selbst- zu den Fremdbeurteilungen anhand der Stellung der Selbstbeurteilung innerhalb der Beurteiltentypen überprüft werden. So ist in folgender Tabelle jeweils derjenige der beiden extrahierten Faktoren wiedergegeben, in dem die Selbstbeurteilung am höchsten lädt. Bei der Abschätzung der Ähnlichkeit zwischen Selbst- und Fremdbeurteilung aufgrund der Ladungsunterschiede muß freilich berücksichtigt werden, daß Ähnlichkeiten nur auf die durch den jeweiligen Faktor erfaßte Varianz zurückgehen. Es sind nur Ladungen bis zur Höhe $\pm .250$ aufgeführt (Tabelle 29).

Arzt	1	2	3	4	5
5)	-.898	3) .702	3) -.693	4) .798	5) -.767
1)	.256	2) -.546	2) .538	3) .468	3) -.574
2)	.229	5) .413	5) .437	5) -.322	1) .267

Tabelle 29: Pro Arzt derjenige rotierte Beurteiltenfaktor, auf dem die Selbstbeurteilung am höchsten lädt

Arzt 1 beurteilt sowohl Arzt 5 als auch sich selbst sehr "spezifisch". Die Ärzte 3, 4 und 5 schätzen sich selbst relativ exponiert ein, ihr Selbst-Urteil kovariiert aber durchaus mit ihrer Beurteilung der Kollegen. Für die Ärzte 2, 4 und 5 scheint Arzt 3 als eine Art Beurteilungsanker fungiert zu haben. Arzt 4 und 5 schätzten sich ihm bezüglich des Explorationsverhaltens ähnlich ein, Arzt 2 beurteilt sich auf der Beurteilerdimension als ihm gegensätzlich. Da Arzt 3 sich selbst auch im Gegensatz zu Arzt 2 sieht und Arzt 1 ein entsprechendes Verhältnis zwischen Arzt 2 und 3 wahrnimmt (dieser Faktor ist in der Tabelle nicht aufgeführt), kann eine gewisse Gegensätzlichkeit im Explorationsverhalten von Arzt 2 und 3 vermutet werden. Zumindest schlägt sich eine solche Gegensätzlichkeit im Urteil der 3 Ärzte nieder.

Die besondere Bedeutung, die Arzt 3 insbesondere im Urteil der Ärzte 2, 4 und 5 genießt, resultiert möglicherweise aus seiner fachlichen Position. Arzt 3 hat von den an der Untersuchung teilnehmenden Ärzten die längste fachliche Erfahrung und war zur Zeit der Untersuchung der einzige habilitierte Arzt.

114. Beurteilungen

Die Beurteilungsfaktorenmatrizen geben Aufschluß darüber, wie sich das durch Diagnosegruppen und Situationen bestimmte Explorationsverhalten laut Urteil der Ärzte ordnet. Es wurden jeweils 3 Faktoren extrahiert.

Die Ähnlichkeit der 5 Faktorenmatrizen ist im Mittel sehr gering (.574), die Streuung mäßig hoch (.096). Damit ordnen - laut Berechnung - die Ärzte die Exploration zu interindividuell nur sehr gering ähnlichen Explorationsstilen.

Anhand der durchschnittlich ähnlichsten (Arzt 5) und unähnlichsten (Arzt 2) Faktorenmatrix soll der Einfluß von Diagnosegruppen und Situationen auf das beurteilte Explora-

tionsverhalten verdeutlicht werden. Die Beurteilungen sind danach gekennzeichnet, bei welcher Diagnosegruppe und zu welcher Exploration sie erfolgten.

Faktor I	Faktor II	Faktor III
16) $-.670$ typisches Explor. Verh.	1) $-.688$ N I	7) $-.595$ N III
12) $-.283$ Z IV	2) $-.584$ N II	5) $.527$ S II
8) $-.255$ S III	3) $-.215$ N III	8) $-.351$ S III
1) $-.252$ N I		13) $.270$ N V
11) $-.243$ S IV		

Tabelle 30: 3 rotierte Explorationsfaktoren des Arztes 5

Faktor I ist gekennzeichnet durch die Beurteilung des durchschnittlichen typischen Explorationsverhaltens, wobei lediglich bei diesem Arzt die Variable typisches Explorationsverhalten diese extrem hohe Ladung erhält. Die weiteren Beurteilungen, die mit dieser einhergehen, sind durch die Kriterien "Psychotische Diagnosegruppen" und "psychiatrisch besonders relevante Explorationssituationen" bestimmt. Faktor I kann somit zusammenfassend als ein Faktor des typisch psychiatrischen Explorationsverhaltens interpretiert werden. Faktor II ist eindeutig gekennzeichnet durch Beurteilungen von Explorationsverhalten gegenüber Neurotikern in psychiatrisch weniger wichtigen Explorationssituationen. Dieser Faktor unterscheidet sich insofern von Faktor I als er durch "normales" Verhalten, d.h. nicht typisch psychiatrisches Explorationsverhalten charakterisiert ist. Faktor III ist im Sinne der Kriterien nicht eindeutig zu interpretieren.

Faktor I	Faktor II	Faktor III
11) $-.490$ S IV	5) $-.567$ S II	2) $-.761$ S I
16) $-.454$ typisches Explor. Verh.	4) $-.456$ N II	1) $-.352$ N I
8) $-.354$ S III	7) $-.406$ N III	3) $-.249$ Z I
15) $-.303$ Z V	3) $-.281$ Z I	
9) $-.267$ Z III	11) $.262$ S IV	

Tabelle 31: 3 rotierte Explorationsfaktoren des Arztes 2

Faktor I entspricht Faktor I des Arztes 5. Er kann wiederum als typisch psychiatrisches Explorationsverhalten interpretiert werden.

Faktor II ist charakterisiert durch das Kriterium Situation, und zwar durch die Situationen I bis III, das Kriterium Diagnosegruppe spielt keine Rolle. Damit ist dieser Faktor sowie Faktor III ein Situationsfaktor. Einfluß auf das Urteil über das Explorationsverhalten nimmt in beiden Fällen die Situation, in der exploriert wird. Faktor I kann bei beiden Ärzten gleich interpretiert werden, nämlich als ein Faktor des typisch psychiatrischen Explorationsverhaltens. Bezüglich der beiden restlichen Faktoren liegen zwischen Arzt 5 und 2 insofern Unterschiede vor, als diese Faktoren bei Arzt 2 vorrangig nur durch das Kriterium Situation bestimmt sind, dagegen bei Arzt 5 durch beide Kriterien, also Situation und Diagnosegruppe. Das Urteil über das Explorationsverhalten orientiert sich demnach bei Arzt 5 sowohl an Situationen, in denen exploriert wird, als auch an Diagnosegruppen, denen die Patienten angehörten. Dabei interagieren beide Kriterien.

Die Interpretation der restlichen 3 Faktorenmatrizen macht deutlich, daß jeweils ebenfalls ein Faktor als "typisch psychiatrisches Explorationsverhalten" interpretierbar ist, während die beiden anderen entweder durch das Kriterium Situation oder durch beide Kriterien bestimmt werden, wobei die Kriterien interagieren.

Interpretativ erweisen sich die 5 Faktorenstrukturen also insofern als inhaltlich ähnlich, als die Faktoren entweder als typisch psychiatrisches Explorationsverhalten interpretiert werden können oder durch die Kriterien Situation und Diagnosegruppe relativ eindeutig charakterisiert sind. Da die die Faktoren kennzeichnenden Situationen bzw. Diagnosegruppen zwischen den Faktorenmatrizen zum Teil variieren, ist die rechnerische ermittelte Faktorenstrukturähnlichkeit gering.

Zusammenfassend kann der Unterschied zwischen dem durch typisches psychiatrisches Explorationsverhalten gekennzeichneten Faktor und den beiden restlichen Faktoren folgendermaßen verdeutlicht werden:

Die Anzahlen der durch die Kriterien Situation und Diagnosegruppe bestimmten Beurteilungen mit einer Ladungshöhe gleich/größer .250 werden getrennt für den Faktor typisches psychiatrisches Explorationsverhalten und die beiden weiteren Faktoren über alle 5 Faktorenmatrizen ermittelt (Tabelle 32).

Faktor	Situation	I	II	III	IV	V	Σ
Typisches psychiatrisches Explorationsverhalten	Diagnosegruppe						
	N	1	1				2
	S	1	1	3	3	2	10
	Z	1		2	3	3	9
	Σ	3	2	5	6	5	
Die beiden restlichen Faktoren	N	5	6	5	3	2	21
	S	2	5	3	1		11
	Z	2		2			4
	Σ	9	11	10	4	2	

Tabelle 32: Zusammenhang zwischen Explorationssituation und Diagnosegruppe für den Faktor "Typisches psychiatrisches Explorationsverhalten" und die beiden restlichen Faktoren

Eine statistische Überprüfung auf Unterschiede ist wegen der Summation über die Faktorenmatrizen der 5 Ärzte nicht gestattet. Eine entsprechende intraindividuelle Überprüfung ist wegen der zu geringen Häufigkeiten nicht möglich. Aus den beiden Tabellen geht jedoch ausreichend deutlich hervor, daß laut Urteil der Ärzte a) typisches Explorationsverhalten die größte Ähnlichkeit mit Explorationsverhalten gegenüber Psychotikern und b) mit Explorationsverhalten in den Situationen III, IV und V besitzt. Unter Ähnlichkeit ist hier die faktorielle Gemeinsamkeit gemeint. Das nicht typische Explorationsverhalten tritt dagegen - laut Urteil der Ärzte - in den Situationen I bis III insbesondere bei neurotischen, aber auch bei schizophrenen Patienten auf. Deutlich wird die Interaktion der beiden, die Beurteilungen definierenden Kriterien pro Tabelle. Die beiden Kriterienstufen "Schizophrene" und "Situation III" nehmen gewissermaßen eine Mittelstellung zwischen den beiden Faktorentypen ein. Sie definieren sowohl das typische psychiatrische Explorationsverhalten als auch das nicht typische. Interessant ist bei diesen Ergebnissen, daß es laut Urteil der Ärzte nur einen typisch psychiatrischen Explorationsstil gibt, während beim nicht typisch psychiatrischen Explorationsverhalten

zwei Explorationsstile unterschieden werden, denn das Explorationsverhalten gegenüber Neurotikern, z.T. auch gegenüber Schizophrenen, wird in den ersten Situationen differenzierter beurteilt. Als Interpretation dafür bietet sich an, daß das typische Explorationsverhalten, das als Verhalten gegenüber Psychotikern gilt, ein normiertes exploratives Verhalten ist. (Eine einzige Norm ist jedoch nicht für alle Ärzte gültig, da im Urteil durchschnittlich zwischen 2 Arzttypen unterschieden wurde). Dagegen unterliegt das Verhalten gegenüber Neurotikern in den ersten Situationen keiner klaren Norm. Laut Urteil kann der explorierende Arzt zwischen zwei Explorationsstilen wählen, um angemessen auf neurotische Patienten einzugehen. Diese beiden Explorationsstile scheinen relativ spezifisch zu sein. Jedenfalls werden sie spezifisch beurteilt. So ist die Anzahl der Beurteilungsdimensionen, die zur Charakterisierung der beiden Explorationsstile verwendet wird, durchschnittlich geringer als die zur Kennzeichnung des typischen Explorationsstiles; die Anzahl der über die insgesamt 5 Kernmatrizen summierten Null-Werte (Werte zwischen $-.999$ und $+.999$) beträgt beim Faktor typisches Explorationsverhalten 6 und bei den restlichen beiden Faktoren 17. Zu erwarten sind statt 17 2 mal 6. Damit wird einerseits die relative Spezifität der Explorationsstile gegenüber Neurotikern in den ersten Situationen belegt, andererseits wird deutlich, daß das Arzt-Polaritäten-Profil gerade zur differenzierten Beurteilung des typischen Explorationsstils geeignet ist.

12. Bestimmung der Beurteilungen durch die individuellen Patienten

Die für jeden Arzt pro Modus durch die Faktoren erfaßten Varianzanteile stimmen recht genau mit den jeweils entsprechenden aus der letzten Darstellung (Abschnitt 1) überein. Die Bestimmung der insgesamt 15 Explorationsbeurteilungen durch die individuellen Patienten führt somit bezüglich der Varianzanteile zu keinen entscheidenden Veränderungen.

Der Vergleich der 5 Skalenfaktorenmatrizen liefert fast gleich hohe Ähnlichkeitskoeffizienten wie unter Abschnitt 1. Dagegen verändert sich die Höhe der Ähnlichkeitskoeffizienten aus dem Vergleich der Beurteiltenfaktorenmatrizen beträchtlich. Unter 1 beträgt der Mittelwert der Koeffizienten $.652$, die Streuung $.126$. Bei diesen Analysen beläuft sich der Mittelwert auf $.469$, die Streuung ebenfalls auf $.126$. Damit werden die Beurteilten zu interindividuell ähnlicheren Typen geordnet, wenn sich die Beurteilung

ihres Explorationsverhaltens an den Kriterien Diagnosegruppe und Situation orientieren. Dagegen wird die Urteiler-Übereinstimmung vermindert, wenn das beurteilte Explorationsverhalten durch die individuellen Patienten bestimmt ist. Gerade umgekehrt ist das Verhältnis bei den beurteilten Explorationsstilen. Der Mittelwert der Ähnlichkeitskoeffizienten steigt von .574 auf .653, und die Streuung erhöht sich ebenfalls (von .096 auf .111), d.h. die beurteilten Explorationen gruppieren sich interindividuell ähnlicher, wenn sie durch die individuellen Patienten statt durch die Kriterien Diagnosegruppe und Situation bestimmt werden.

Interpretiert man höhere Interurteilerübereinstimmung als Hinweis auf eine einfache und sichere Beurteilung, dann erleichtern die beiden Kriterien der Explorationen die Typisierung der Ärzte aufgrund der Exploration, während die Typisierung der Exploration selbst zu Explorationsstilen leichter gelingt, wenn die Explorationen durch die individuellen Patienten charakterisiert sind. Daraus kann die weitere Schlußfolgerung gezogen werden, daß bei der Beurteilung der Ärzte vorrangig auf das Explorationsverhalten in bestimmten Explorationssituationen und gegenüber bestimmten Patientengruppen Wert gelegt wird, während auf die Einschätzung der Exploration selbst der jeweils explorierte Patient ausschlaggebend Einfluß nimmt.

Die Analyse der Beurteilungsfaktoren, die Aufschluß über Explorationsstile geben, liefert ein ähnliches Bild wie die entsprechenden Analysen im Abschnitt 1. Es tritt wiederum jeweils ein Faktor auf, der durch typisches Explorationsverhalten und die Exploration von Schizophrenen und Cyclothymen gekennzeichnet ist, während die jeweils beiden restlichen Faktoren durch Explorationen von Neurotikern z.T. auch durch Explorationen von Schizophrenen näher bestimmt sind.

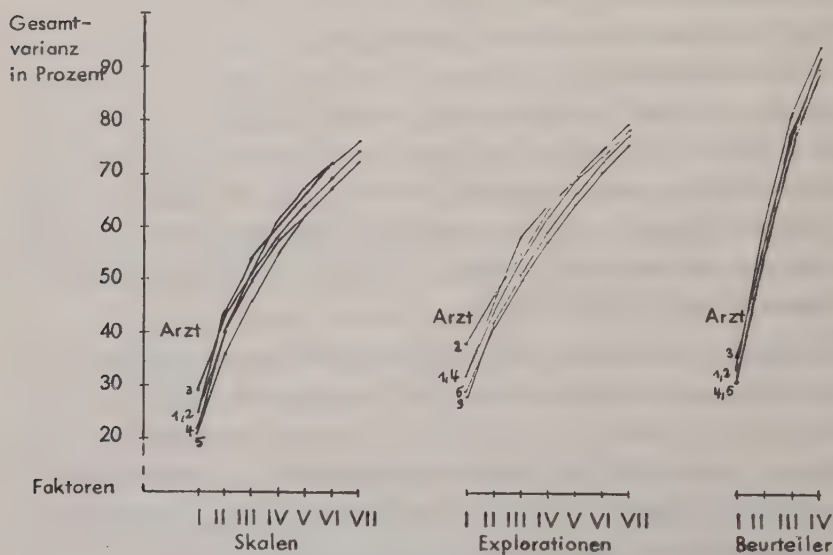
2. Ärzte als Beurteilte

Bei diesem rechnerischen Ansatz gehen alle Ärzte als Beurteiler bei der Beurteilung jeweils eines Arztes in die 3 Modi-FA ein. Errechnet werden somit pro beurteiltem Arzt durchschnittliche Urteilsstrukturen, durchschnittliche Explorationsstile und außerdem Beurteilertypen. Von besonderem Interesse ist hier die inhaltliche Bedeutung der durchschnittlichen und daher vermutlich relativ ausgewogenen Beurteilungsdimensionen, deren Abhängigkeit vom jeweils Beurteilten, d.h. der spezifischen Information, und weiterhin die Anzahl der pro beurteiltem Arzt wahrgenommenen Explorationsstile sowie

die jeweilige Anzahl von Beurteilertypen. Die Beurteilten sind durch die Bedingungen Diagnosegruppe und Situationen bestimmt.

21. Differenziertheit der Urteile

Die folgende graphische Darstellung 11 gibt pro beurteiltem Arzt für jeden Modus die kumulierten Varianzanteile der unrotierten Faktoren wieder.



Graph. Darst. 11: Kumulierte prozentuale Gesamtvarianzanteile der unrotierten Faktoren pro Modus und beurteiltem Arzt

Bei den beurteilten Ärzten 1 und 2 können jeweils drei Skalenfaktoren, bei den Ärzten 3, 4 und 5 jeweils 4 Skalenfaktoren als bedeutsam gelten, jedoch sind die Unterschiede in Höhe und Verlauf der Eigenwerte so gering, daß die Entscheidung für 3 bzw. 4 bedeutsame Skalenfaktoren verhältnismäßig willkürlich ist.

Lediglich zwischen den Varianzanteilen des jeweils 1. Faktors treten etwas größere Diskrepanzen auf. So wird z.B. Arzt 3 - insbesondere im Verhältnis zu Arzt 5 - mit einem relativ varianzstarken 1. Faktor beurteilt. Insgesamt sind jedoch die Unterschiede der Differenziertheit der Beurteilung der einzelnen Ärzte mäßig groß. Damit nimmt die Information, d.h. der einzelne beurteilte Arzt, nur mäßigen Einfluß auf den Differenziertheitsgrad der Skalenbenutzung.

Etwas größer sind die Differenziertheitsunterschiede bei der Beurteilung der 15 bzw. einschließlich des typischen Explorationsverhaltens 16 Explorationen. So wird z.B. bei Arzt 4 deutlich zwischen 3 Explorationsstilen differenziert, während Arzt 2 in etwa als nur einen, relativ globalen Explorationsstil verwendend wahrgenommen wird.

Der Varianzanteil des jeweils 1. Faktors der Beurteilmatrizen ist hier eindeutig geringer als bei jeder der Patientenbeurteilungen. Eindeutig bedeutsam sind entsprechend jeweils mindestens 2 Faktoren. Die Ärzte differieren damit stärker bei der Beurteilung von Kollegen (einschließlich sich selbst) als bei der Beurteilung von Patienten. Als Erklärung dafür kann wiederum die gemeinsame Geübtheit in der Beurteilung von Patienten herangezogen werden. Dieses gemeinsame Training fehlt bei der Einschätzung von Kollegen, so daß es zu größeren Unterschieden zwischen den Beurteilten kommt.

22. Skalen

Es werden pro beurteiltem Arzt 4 Skalenfaktoren extrahiert und die übliche rechnerische Weiterverarbeitung vorgenommen.

Der mittlere Ähnlichkeitskoeffizient zwischen den 5 Skalenfaktorenmatrizen beträgt .714, die Streuung .057. Die Skalenfaktorenstrukturen weisen damit einerseits deutliche Ähnlichkeiten auf, andererseits tritt jedoch der Unterschied in Abhängigkeit vom Beurteilten, d.h. von der spezifischen Information klar hervor. Dieses Ergebnis widerspricht zwar einerseits verschiedenen Untersuchungen (z.B. NORMAN und GOLDBERG, 1967), in denen die relative Unabhängigkeit der Urteilsstrukturen von den Beurteilten nachgewiesen wurde, entspricht andererseits jedoch den Untersuchungen von WARR und KNAPPER (1968) und insbesondere der diagnostischen Forderung nach Anwendung spezifischer, für den jeweils Beurteilten besonders charakteristischer Urteilsdimensionen.

Dargestellt werden sollen hier die Skalenfaktoren bei der Einschätzung von Arzt 4. Diese Faktorenmatrix zeigte die durchschnittliche zweitgrößte Übereinstimmung zu den anderen entsprechenden Faktorenmatrizen.

Die Darstellung der Faktoren erfolgt wie üblich (Tabelle 33).

Faktor

I	13)	-.495	originell vorgehend	-	schematisch vorgehend
	5)	-.455	im Verhalten ausdrucksvoll	-	im Verhalten farblos
	20)	.298	gelöst	-	gespannt
	10)	.276	lebhaft	-	still
	1)	-.274	aktiv	-	passiv
II	2)	-.549	rational	-	gefühlsmäßig
	6)	.527	zielgerichtet	-	weitschweifig
	21)	.342	gleichbleibend	-	wechselnd
	20)	.305	gelöst	-	gespannt
III	18)	-.659	beharrend	-	schnell aufgebend
	16)	-.402	kontrolliert	-	unbefangen
	11)	.355	mißtrauisch	-	arglos
	19)	-.329	interessiert	-	gelangweilt
	12)	-.254	streng	-	gewährend
IV	10)	-.462	still	-	lebhaft
	7)	.428	selbstkritisch	-	selbstgefällig
	14)	.389	sachlich	-	selbstbezogen
	4)	-.359	geduldig	-	ungeduldig

Tabelle 33: 4 rotierte Skalenfaktoren bei der Beurteilung von Arzt 4

Faktor I ist gekennzeichnet durch die persönlichkeitspezifische Eigenschaftsdimension "Dynamik", mit der ein "origineller (vs. schematischer)" Explorationsstil einhergeht. Wichtigstes Charakteristikum von Faktor II ist "Zielgerichtetheit" in der Exploration, während Faktor III als "Engagement und Ausdauer" beim Explorieren interpretiert werden kann. Faktor IV ist durch "Ruhe und Problemzentriertheit" beim Explorieren charakterisiert.

Wie schon erwähnt, spielen bewertende Aspekte bei der Beurteilung des Explorationsverhaltens von Kollegen sicher eine nicht unbedeutende Rolle. Bei den hier dargestellten 4 Faktoren dürfte jeweils der linke Pol der Eigenschaftsskalen eine positive Bewertung, der rechte Pol entsprechend eine negative Bewertung implizieren.

23. Beurteilungen

Anhand der Varianzanteile der Faktoren aus den 5 Faktorenmatrizen waren deutliche Unterschiede im Grad der Differenziertheit der beurteilten Explorationsstile nachgewiesen worden. Auch die Faktorenstrukturen unterscheiden sich erheblich; in die Ähnlichkeitsvergleiche gingen nach genauer Analyse der Eigenwerte jeweils 3 Faktoren ein.

Der mittlere Ähnlichkeitskoeffizient beträgt .545, die Streuung der Koeffizienten .057.

Die Interpretation der Faktoren führt zu demselben Ergebnis wie unter "Ärzte als Beurteiler". Jeweils ein Faktor ist bestimmt durch die typische Exploration (Variable 16) und durch Explorationen von Psychotikern in psychiatrisch relevanten Explorationssituationen. Die beiden anderen Faktoren sind gekennzeichnet durch die Explorationen von Neurotikern und auch Schizophrenen in den ersten 3 Explorationssituationen.

Die entsprechende Auszählung der Beurteilungen mit Ladungshöhen gleich/größer .250 resultiert in einem ähnlichen Ergebnis wie unter "Ärzte als Beurteiler" (Tabelle 34).

Faktor	Situation	I	II	III	IV	V	Σ
Typisches psychiatrisches Explorationsverhalten	Diagnose- gruppe						
	N	1		1	1	2	5
	S	1	1	3	2	3	10
	Z		2	2	1	3	8
	Σ	2	3	6	4	8	

Die beiden rest- lichen Faktoren	Situation	I	II	III	IV	V	Σ
	Diagnose- gruppe						
	N	5	4	5	4	4	22
	S	2	4	3	2	1	12
	Z	5	1	1	2	2	11
	Σ	12	9	9	8	7	

Tabelle 34: Zusammenhang zwischen Explorationssituation und Diagnosegruppe für den Faktor "Typisches psychiatrisches Explorationsverhalten" und die beiden restlichen Faktoren

24. Beurteiler

Es wurden jeweils 2 Faktoren extrahiert, zur Interpretation rotiert und die 5 Faktorenmatrizen auf ihre Ähnlichkeit hin verglichen. Anhand der Ähnlichkeitskoeffizienten kann überprüft werden, ob und wie weit die Gruppierung der Beurteiler in Abhängigkeit von den jeweils Beurteilten steht. Niedrige Koeffizienten deuten einen starken Einfluß der Beurteilten auf die Beurteilertypen an.

Das Ergebnis der Ähnlichkeitsvergleiche ist insofern besonders interessant, als die Streuung, bei einem Mittelwert der Koeffizienten von .748 sehr groß ist, nämlich .144. Demnach ist eine eindeutige Abhängigkeit der Beurteilertypen von den Beurteilten in einigen Fällen gegeben, in anderen Fällen veranlassen Beurteilte fast gleiche Beurteilertypen (Tabelle 35).

	2	3	4	5
1	.577	.832	.729	.702
2		.568	.932	.931
3			.654	.640
4				.917

Tabelle 35: Ähnlichkeitskoeffizienten aus dem Vergleich der 5 Beurteilerfaktorenmatrizen

Insbesondere die Ärzte 2, 4 und 5 rufen bei den Beurteilern jeweils gleiche Urteilstendenzen hervor, so daß es zu einer jeweils sehr ähnlichen Gruppierung der Beurteiler kommt. Diese hohe Übereinstimmung der Beurteilertypen kann als Hinweis auf Klarheit des Explorationsverhaltens der 3 Ärzte und entsprechend als Hinweis auf Einfachheit der Beurteilbarkeit interpretiert werden; die Beurteiler können in diesen 3 Fällen ihrem üblichen Urteilsstil nachkommen, und folglich ist das Verhältnis der Beurteiler zueinander konstant.

Im Gegensatz zum Explorationsverhalten der Ärzte 2, 4 und 5 scheint das der Ärzte 1 und 3 weniger transparent zu sein. Entsprechend ändert sich die jeweilige interindividuelle Übereinstimmung der Beurteiler. Die Beurteilung erweist sich als schwierig, der übliche Urteilsstil muß aufgegeben werden, und folglich kommt es zu einer jeweiligen Neugruppierung der Beurteiler.

Auf die weitere Berechnung, bei der die Beurteilungen durch die individuellen Patienten bestimmt sind, wurde verzichtet.

E ZUSAMMENFASSUNG, DISKUSSION, KRITIK

Ziel einer jeden Diagnostik ist es, Patienten zu kategorisieren und damit zu beschreiben. Die Psychiatrie stellt zu diesem Zweck ein diagnostisches System zur Verfügung, das formal aus konjunktiven, insbesondere aber aus disjunktiven Klassen besteht. Die Psychologie lieferte bisher kein ausreichend umfassendes, eindeutig definiertes nosologisches System. Diese Tatsache hat zur Folge, daß jeder psychologisch-diagnostisch Arbeitende ein eigenes psychologisch-nosologisches System anwendet. Dieses System mit meist relationalen Klassen und den sie definierenden Beurteilungsdimensionen ist seinem Anwender gewöhnlich nicht ausreichend bewußt. Solche Systeme lassen sich jedoch - wie z.B. in dieser Untersuchung - ermitteln. Dazu ist es notwendig, daß der Urteiler eine größere Anzahl von Patienten anhand von Skalen einstuft. Diesem Prozeß, aber auch dem psychiatrisch-diagnostischen Prozeß, der zur Kategorisierung der Patienten in vorgegebene Diagnosegruppen führt, liegt ein konvergenter Denkvorang zugrunde, der dadurch relativ kompliziert ist, daß die zur Kategorisierung erforderliche Information häufig - insbesondere bei der psychologischen Beurteilung - einer interpretativen Aufbereitung bedarf.

In beiden Fällen führt der Denkvorgang als Entscheidungsprozeß zu einer näheren Charakterisierung der Patienten. Beim psychologischen Urteil gilt das sowohl für die unmittelbare Einstufung von Patienten anhand der Skalen als natürlich auch nach Ermittlung von Patientenklassen, denen die Patienten zugehören.

Aufgrund der Charakterisierung verlieren die Patienten zwar ihre Individualität, sie werden jedoch manipulierbar, d.h. es ergeben sich unmittelbar Handlungskonsequenzen. Diese Handlungskonsequenzen bestehen nach einer psychiatrischen Diagnose in Therapieanweisungen, die psychiatrische Diagnose hat vorrangig diese Funktion. Entsprechende Handlungskonsequenzen kann ein von Psychiatern getroffenes psychologisches Urteil nicht nach sich ziehen, da ein psychologisches Diagnosesystem mit Therapieanweisungen fehlt. Vielmehr dürfte das psychologische Urteil zweierlei Funktion haben. Einerseits unterstützt es vermutlich das psychiatrische Urteil - insbesondere bei der Diagnose von Neurosen -, andererseits dürfte es jedoch insofern eine dem psychiatrischen Urteil äquivalente eigenständige Funktion besitzen, als es zu einer psychologischen Charakterisierung von Patienten führt, die dem diagnostizierenden Psychiater ein erstes Konzept über die Patienten liefert und - im Sinne der Handlungskonsequenz - ein zielgerichtetes Eigenverhalten des Psychiaters gestattet. Diese Funktion hat das psychologische Urteil in allen sozialen Interaktionen, in denen sich die Akteure relativ schnell, für sie kognitiv wenig bewußt, häufig vorurteilshaft voneinander ein Bild machen, um gerichtet handeln zu können.

Aus diesen Überlegungen über die unterschiedliche Funktion des psychologischen und psychiatrischen Urteils folgt die anfangs formulierte psychologische Basis-Hypothese, das psychologische Urteil impliziere einen Denkvorgang, der - im Gegensatz zum psychiatrischen Urteil, dem ein bewußtes Abtesten von Symptomen zugrunde liegt - auf einer niederen kognitiven Bewußtseinsstufe verlaufe, sich folglich relativ schnell stabilisiere und von bestimmten Patientengruppen verhältnismäßig wenig differentiell beeinflusst werde.

Grob zusammengefaßt konnte diese Vermutung bestätigt werden.

Im einzelnen kamen folgende Ergebnisse zustande, wobei hier nur die wichtigsten aufgeführt werden sollen.

Entsprechend einer der eingangs gestellten Hypothesen sind für den Durchschnitt der Urteiler die Beurteilungsdimensionen als kognitiver Beurteilungsraum inhaltlich nur sehr wenig von Informationsausmaß und -inhalt abhängig, d.h. die Beurteilungsskalen

werden sehr reliabel interpretiert. Entgegen der auf Untersuchungen (KOLTUV, 1962; COHEN, 1969) und den Überlegungen zum Halo-Effekt und logischen Fehler basierenden Hypothese konnte eine beständige Ausdifferenzierung der Beurteilungsdimensionen bei zunehmender und psychiatrisch wichtiger werdender Information nicht nachgewiesen werden.

Beide in dieser Untersuchung gewonnenen Ergebnisse bestätigen das Modell von SARBIN et al. (1960) und machen deutlich, daß die Psychiater über außerordentlich stabile, damit gleichzeitig aber auch rigide Beurteilungsdimensionen zur Beurteilung von Patienten verfügen.

Informationsausmaß und -inhalt haben nicht nur auf die Urteilsdimensionen minimalen Einfluß, sondern auch auf die Mittelwerte der einzelnen Skalen des Polaritäten-Profiles. So ergab die intraindividuelle Überprüfung pro Skala auf Unterschiede in der zentralen Tendenz zwischen den drei Urteilen in den 105 Fällen nur 7 signifikante Differenzen (Friedman Test; in 7 Fällen $p < 0,05$).

Die innerhalb des konstanten kognitiven Beurteilungsraums lokalisierten Patiententypen als fast ausschließlich relationale Klassen sind gegenüber Informationsausmaß und -inhalt weniger resistent. Sie differenzieren sich insbesondere von Urteil 1 nach 2 innerhalb des Beurteilungsraumes aus und ändern sich auch inhaltlich. Zusätzliche und psychiatrisch besonders wichtige Information aus den Situationen IV und V trägt dann nur noch vorrangig zu einer Veränderung der Differenziertheit bei (Urteil 3). Mit diesen Ergebnissen werden die Hypothesen bezüglich Inhalt und Differenziertheit der Patiententypen bestätigt.

Die den kognitiven Raum aufspannenden Beurteilungsdimensionen sind relativ eindeutig interpretierbar und zeigen Zusammenhänge zu psychopathologischen Syndromen. Nur eine der 4 Beurteilungsdimensionen impliziert Bewertungsaspekte. Die bewertende Stellungnahme richtet sich weniger gegen die Patienten als soziale Partner selbst als vielmehr gegen Eigenschaften der Patienten (stumpf, undifferenziert), die sie für den explorierenden, eine Diagnose anzielenden Psychiater uninteressant erscheinen lassen.

Die Patiententypen sind mit Hilfe der als Außenkriterium fungierenden Diagnosen in allen 3 Urteilen nur schwer interpretierbar, womit die Hypothese der geringen Übereinstimmung zwischen Patiententypen aus dem psychologischen Urteil und psychiatrischen Patientenklassen nachgewiesen ist. Eindeutig ist lediglich ein unipolarer Schizo-

phrenie-Faktor. Demnach sind Schizophrene auch die einzigen Patienten, die laut psychologischem Urteil noch am ehesten einer kategorialen Klasse angehören (vgl. auch die Ergebnisse zur Kernmatrix). Dieses Ergebnis ist plausibel, da sich Schizophrene in ihrer Symptomatik stärker von anderen Patientengruppen unterscheiden als jede andere Patientengruppe.

Ein direkter Vergleich zwischen dem Verlauf des psychologischen und psychiatrischen Urteils muß die Patientenklassen betreffen, da beim psychiatrischen Urteil nur die Kategorisierung in solche Klassen erfaßt wird. Diese kategorialen Klassen, die der Psychiater als diagnostische Hypothesen an die Patienten heranträgt, sind mit ihren sie definierenden Symptomen hypothetische Konstrukte. Es kann davon ausgegangen werden, daß diese Klassen von Informationsausmaß und -inhalt über bestimmte Patienten unabhängig sind. Eine Veränderung dürfte nur durch Anweisung anderer oder eigene Weiterbildung hervorgerufen werden. Eine solche Situation war während der jeweiligen Untersuchung eines Patienten nicht gegeben. Aufgrund dieser Überlegungen kann der Vergleich der Patientenklassen aus psychologischem und psychiatrischem Urteil spezifiziert werden:

In beiden Fällen sind die die Klassen bestimmenden Merkmale konstant (siehe Ergebnis zu Beurteilungsdimensionen aus psychologischem Urteil). Somit projizieren die Psychiater sowohl beim psychologischen als auch beim psychiatrischen Urteil konstante Merkmale (Dimensionen im psychologischen und Symptome im psychiatrischen Urteil) auf die zu klassifizierenden Patienten. Daher ist der die Patientengruppen betreffende einzig mögliche Verlaufsvergleich zwischen psychologischem und psychiatrischem Urteil äußerst sinnvoll, da eine Veränderung im Urteilsverlauf auch nur die Patientenklassen betrifft. Laut Untersuchungsergebnis stabilisieren sich entsprechend der Hypothese die Patientenklassen aus dem psychologischen Urteil schneller als die aus dem psychiatrischen Urteil, d.h. die psychologische Sicht der Patienten stabilisiert sich schneller als die psychiatrische. Diese Stabilisierung gilt insbesondere für den Inhalt, weniger für die Differenziertheit der Patiententypen. Im psychiatrischen Urteil kann zwischen diesen beiden Aspekten deutlich unterschieden werden. So betrifft die Stabilität, die zwischen 2. und 3. Urteil eher geringer ist als zwischen 1. und 2., den Inhalt der Patientengruppen, d.h. die psychiatrische Sicht der Patienten.

Dagegen sind die Beschreibungen als Ausmaß für die undifferenzierte Beurteilung der Patienten interpretierbar; sie fallen gleichmäßig von 1. bis 3. Urteil. Beschreibungen - bisher meist interpretiert als Urteilsunsicherheit bzw. Beurteilungsmehrdeutigkeit - sind

nämlich gleichermaßen Anzeichen dafür, daß eine differenzierte Kategorisierung in eine der Klassen des nosologischen Systems nicht gelungen ist; die Patienten erscheinen mehrdeutig und damit im Sinne des nosologischen Systems undifferenziert.

Im psychologischen Urteil ist diese klare Trennung zwischen Differenziertheit und Inhalt der Patiententypen nicht möglich, da in die Ähnlichkeitsvergleiche der Faktorenmatrizen mit jeweils 4 Faktoren beide Aspekte eingehen. Im 1. Urteil treten nämlich nur 3 Patiententypen auf, was eine Verminderung der entsprechenden Ähnlichkeitsvergleiche impliziert. Daher kann - wie schon ausgeführt - nur die Aussage getroffen werden, daß sich die Patiententypen aus dem psychologischen Urteil ab 2. Urteil vorrangig bezüglich des Inhalts der Patiententypen stabilisieren, sich aber zum 3. Urteil weiterhin ausdifferenzieren.

Im psychiatrischen Urteil findet dagegen eine stetige Veränderung der beiden trennbaren Aspekte der Patientenklassen (psychiatrische Sicht der Patienten) von Urteil 1 bis 3 statt. Die stetige Veränderung des psychiatrischen Urteils ist auch anhand des gleichmäßig zunehmenden Anteils richtiger Diagnosen nachweisbar.

Hinweise auf den relativ geringen Effekt von Informationsausmaß und -inhalt auf das psychologische Urteil liefert auch die beurteilte Stereotypievielfalt der Patienten. So konnte die Hypothese, daß bei zunehmender und psychiatrisch wichtiger werdender Information die Patienten in ihrer Symptomatik klarer erfaßt und entsprechend mit größerer stereotyper Symptomvielfalt beurteilt werden, nicht bestätigt werden. Allgemein hat sich das psychologische Urteil durchschnittlich nur wenig am psychologischen Bild typischer Patienten orientiert. Jedoch erfolgte die Beurteilung einzelner Patienten durch einen Arzt häufig eindeutig entsprechend der Vorstellung von einem typischen Patienten.

Zusammenfassend kann festgestellt werden, daß die psychologische Sicht der Patienten weniger abhängig von Informationsausmaß und -inhalt ist als die psychiatrische Sicht. Das psychologische Urteil zeigt sich jedoch nicht nur von dieser Bedingung relativ unabhängig. Auch das spezifische Stimulusmaterial, nämlich - laut Außenkriterium - die drei Patientengruppen, üben nur minimalen Einfluß auf das psychologische Urteil aus, wofür auch die geringe Übereinstimmung der Patiententypen aus psychologischem Urteil mit den psychiatrischen Diagnosen der Patienten als Hinweis gelten kann. Weiterhin differieren die drei Patientengruppen pro Urteil nicht signifikant bezüglich ihrer stereotypen Beurteilbarkeit. Und desgleichen unterscheidet sich die psychologische Beurteilbarkeit (Kommunalität) der drei Patientengruppen pro Urteil nicht signifikant.

Dagegen treten - wie erwartet - bezüglich der psychiatrischen Beurteilbarkeit der drei Patientengruppen im Sinne der schweren bzw. leichten Beurteilbarkeit immerhin im 1. Urteil signifikante Unterschiede auf, und zwar sind Schizophrene - übereinstimmend mit den Ergebnissen aus der Literatur - leichter, und somit valider beurteilbar als Neurotiker und insbesondere Cyclothyme. Zunehmende und psychiatrisch wichtiger werdende Information führt dann zu einem Ausgleich der psychiatrischen Beurteilbarkeit.

Weiterhin ließ sich nachweisen, daß Informationsausmaß und -inhalt differentiellen Einfluß auf das psychiatrische Urteil pro Diagnosegruppe nehmen. Dieser Effekt konnte sowohl anhand der insgesamt abgegebenen als auch anhand der jeweils richtigen Diagnosen aufgezeigt werden. Während die minimale Information aus Situation I Diagnosetendenzen als Angleichung der Anzahl abgegebener Diagnosen an die Basisraten nicht verhindern konnte, führte dann die Information aus Situation II und III zu einem starken Anstieg der Anzahl von Neurose-Diagnosen, während die zusätzliche Information aus Situation IV und V eine relativ starke Erhöhung der Anzahl von Schizophrenie- und Cyclothymie-Diagnosen zur Folge hatte. Im dritten Urteil waren die Diagnosetendenzen fast völlig ausgeglichen.

Entsprechend wuchs die Anzahl richtiger Neurose-Diagnosen insbesondere zum zweiten Urteil und die richtiger Schizophrenie- und Cyclothymie-Diagnosen vor allem zum dritten Urteil.

Bei der Interpretation dieses Ergebnisses scheint eine Trennung zwischen Informationsausmaß und -inhalt gerechtfertigt; sehr wahrscheinlich bewirkt nämlich der Informationsinhalt diesen differentiellen Effekt. Demnach scheint für psychiatrische Diagnosen inhaltlich spezifische Information erforderlich zu sein. Daraus kann gefolgert werden, daß - entsprechend den Ausführungen in der Einleitung - hauptsächlich eine Strategie der positiven Klassifikation eingeschlagen wird. Das heißt, die Information muß Aufschluß darüber geben, ob bestimmte Symptome vorliegen, das Nicht-Vorhandensein von Symptomen scheint nicht ausreichend zur Klassifikation beizutragen (vgl. BRUNER et al., 1956, die unter anderen Bedingungen und an anderem Material ähnliche Effekte nachwiesen). Die Information aus Situation II und insbesondere III läßt entsprechende Schlußfolgerungen in Bezug auf neurotische Symptomatik, die Information aus Situation IV und V auf psychotische Symptomatik zu; die Information nimmt Einfluß auf die psychiatrische Beurteilbarkeit der Patienten. Es scheint sogar so zu sein, daß die Information zur positiven Klassifikation verleitet, wie die starke Erhöhung der Anzahl von

(falschen) Neurose-Diagnosen im 2. Urteil und die Vermehrung der (falschen) Schizophrenie- und Cyclothymie-Diagnosen im 3. Urteil vermuten lassen. Dieses Ergebnis der einerseits nicht angemessenen, andererseits nicht optimalen Ausnutzung bestimmter Informationsinhalte ist von praktischer Bedeutung. Trainingsprogramme zum Erlernen von Exploration sollten diese Ergebnishinweise berücksichtigen.

Ähnliche Ergebnisse zeigen sich im psychologischen Urteil nicht. Die Kommunalitäten als Ausdruck für die Urteilsbestimmtheit unterscheiden sich in ihrer zentralen Tendenz pro Diagnosegruppe zwischen den 3 Urteilen nicht signifikant. Die geringen χ^2 -Werte weisen darauf hin, daß praktisch keinerlei Veränderung der differentiellen Urteilsbestimmtheit stattfindet.

Der hier versuchte Vergleich zwischen psychologischem und psychiatrischem Urteil bezüglich ihrer Abhängigkeit von Informationsausmaß und -inhalt sowie von den drei Patientengruppen ist schwierig, da er interpretativ erfolgt. Eine statistische Überprüfung auf Unterschiede ist nicht möglich.

Es erscheint jedoch aufgrund der Ergebnisse gerechtfertigt festzustellen, daß das psychologische Urteil von den beiden Bedingungen weniger abhängig ist als das psychiatrische Urteil; oder wiederum klarer formuliert: die psychologische Sicht der Patienten (Patiententypen) ist durch die beiden Bedingungen weniger beeinflussbar als die psychiatrische Sicht der Patienten (Anzahl von Patienten in Diagnosegruppen). Insbesondere ist die frühere Stabilisierung der psychologischen Sicht der Patienten hervorzuheben.

Die verschiedenen Ergebnisse wurden z.T. vorhergesagt. Einige der Hypothesen basieren auf den Ergebnissen anderer Untersuchungen, die den Unterschied zwischen psychologischem und psychiatrischem Urteil betreffenden Hypothesen dagegen auf der Überlegung, daß das psychologische Urteil anhand von Skalen auf einer niederen Bewußtseisebene abläuft als das psychiatrische Urteil. So ist insbesondere die Interpretation der Information außerordentlich stabil, wie die konstanten Urteilsdimensionen belegen, d.h. der Schluß von Information auf Eigenschaften (Skalen) ist als sehr eingeschliffen zu betrachten. Als bis zu einem bestimmten Grad abhängig von Informationsausmaß und -inhalt erweist sich die psychologische Sicht der Patienten, während dagegen die Differenziertheit der Patiententypen deutlich mit Informationsausmaß und -inhalt variiert, aber nicht zu einer qualitativ anderen Sicht der Patienten führt. Das psychologische Urteil über die Patienten scheint somit - wie eingangs ausgeführt - die übliche Funktion innerhalb einer sozialen Interaktion zu haben, d.h. es liefert sehr schnell ein stabiles Bild vom

Patienten, das lediglich durch Ausdifferenzierung klarer wird. Es steckt den Rahmen für eine exakte symptombezogene psychiatrische Analyse ab.

Aus dieser Überlegung resultiert Kritik an Untersuchungen wie dieser, in denen das psychologische Urteil durch Einschätzen anhand von Skalen abgegeben wird. Eine Alternative wäre, das psychologische Urteil in ähnlicher Form abgeben zu lassen wie das psychiatrische, d.h. auch das psychologische Urteil sollte in Angabe psychologischer Klassenzugehörigkeit der Patienten, möglichst mit spezifischen Beschreibungen, erfolgen. Andererseits kann man jedoch davon ausgehen, daß der Psychiater bei der praktischen psychologischen Beurteilung eines Patienten nicht eine Kategorisierung in eine psychologische Klasse vornimmt, zumal ausreichend spezifizierte Klassen fehlen. Es ist zu vermuten, daß er auch in der praktischen Arbeit bestimmte konstante psychologische Dimensionen an die Patienten heranträgt und dann - etwa wie in dieser Untersuchung - sich schnell ein stabiles Bild vom Patienten macht.

Es ist an dieser Stelle angebracht, die in dieser Untersuchung aufgewiesene Funktion des psychologischen Urteils und die damit verbundene schnelle Stabilisierung dieses Urteils einer kritischen Betrachtung zu unterziehen. In der Einleitung wurde auf die Mängel des klassischen psychiatrischen nosologischen Systems mit seinen logischen Klassen hingewiesen und ein psychologisches Diagnose-System mit relationalen Klassen, das dem realen Verhalten der Patienten besser Rechnung trägt, als zunehmend wichtiger werdende Alternative aufgezeigt. Bei Anwendung eines solchen psychologischen Systems wird sich die Funktion des psychologischen Urteils auch für den dann psychologisch urteilenden Psychiater ändern müssen. Das impliziert, daß neue Information in weit stärkerem Maße nach psychologisch bedeutsamen Inhalten abgesucht werden wird und somit das Urteil - vermutlich ähnlich wie üblicherweise beim Psychologen - flexibler sein wird.

In der bisherigen Diskussion wurde das durchschnittliche psychiatrische und psychologische Urteil behandelt. Der bisher entstandene Eindruck des psychiatrischen und psychologischen Urteilens ist insofern unrealistisch als - entsprechend der Erwartung - beträchtliche interindividuelle Unterschiede zwischen den Psychiatern bestehen. Im psychologischen Urteil verstärken sich diese interindividuellen Differenzen bezüglich Inhalt und Differenziertheit der Faktoren erheblich, wenn das Urteilen ungeübt ist wie die Beurteilung explorativen Verhaltens von Psychiatern. Aus der größeren Labilität des ungeübten psychologischen Urteilens resultiert sogar eine deutliche Abhängigkeit

von Urteilsstrukturen, beurteilten Explorationsstilen sowie sogar Beurteilertypen vom Stimulusmaterial, also den Beurteilten.

Bei der psychologischen Beurteilung der Patienten liegen deutliche interindividuelle Unterschiede in Bezug auf Inhalt und Differenziertheit der Faktoren vor. Die Differenzen gehen insbesondere auf die unterschiedliche Sicht der Patienten, weniger auf unterschiedliche Urteilsstrukturen zurück. Diese interindividuellen Differenzen sind jedoch nicht übermäßig groß, wie der jeweils nur eine Beurteilerfaktor belegt. Zunehmende Information trägt zu einer Verminderung der Unterschiede bei. Eindeutig sind die interindividuellen Differenzen auch anhand der Skalenmittelwerte des Patienten-Polaritäten-Profiles nachweisbar. So zeigen sich in 32 der 63 Fälle signifikante interindividuelle Unterschiede in der zentralen Tendenz (Friedman Test; in 32 Fällen $p < 0,05$).

Während sich die interindividuellen Unterschiede im psychologischen Urteil mit zunehmender und psychiatrisch relevanter werdender Information vermindern (siehe Varianzanteile der Beurteilerfaktoren), tritt im psychiatrischen Urteil ein gegenläufiger Trend auf. Zwar ist im 1. Urteil der interindividuelle Unterschied bezüglich der Anzahl der B signifikant, dieser Unterschied verschwindet jedoch schon im 2. Urteil, und im 3. Urteil differieren die Ärzte signifikant in ihrer Urteilsfähigkeit. Die globale Hypothese, die Ärzte unterscheiden sich deutlich in der Validität ihrer Urteile, kann somit nicht völlig bestätigt werden. Die entsprechende Hypothese bezüglich der Stabilität der Urteile erweist sich sogar als falsch. Dieses Ergebnis steht im Widerspruch zu Angaben in der Literatur. Der Widerspruch läßt sich insofern erklären, als der Zeitabstand zwischen den Urteilen in dieser Untersuchung weitaus geringer war als die in den Untersuchungen genannten Zeitabstände, so daß interindividuelle Unterschiede - etwa wie bei der Validität der Urteile - nicht ausreichend deutlich hervortreten konnten. Daß der Zeitabstand zwischen den Urteilen entscheidend Einfluß auf die Ergebnisse genommen hat, zeigt der tendenzielle interindividuelle Unterschied bei Stabilität 1/3.

Das Ergebnis, die Psychiater differieren in ihrer Urteilsfähigkeit erst im 3. Urteil signifikant, bestätigt indirekt den Befund, daß die kognitiven Prozesse beim psychiatrischen und psychologischen Urteilen unterschiedlich sind. Im psychiatrischen Urteil wird die Information in einem, auf einer relativ hohen Bewußtseinssebene ablaufenden kognitiven Prozeß zur Überprüfung von Hypothesen verwendet. Entsprechend kristallisieren sich unterschiedliche psychiatrisch-diagnostische, geschulte Fertigkeiten erst nach Verarbei-

tung von viel und psychiatrisch relevanter Information heraus. Dagegen werden im psychologischen Urteil nicht direkt Hypothesen überprüft, und demgemäß spielen diagnostische, geschulte Fertigkeiten eine mindere Rolle; da Ausbildungsunterschiede nebensächlich sind und somit die Psychiater von gleichen fachlichen Voraussetzungen ausgehen, führt folglich zunehmende Information zu einem Ausgleich von Unterschieden.

Die bisherigen Aussagen bedürfen jedoch einer sehr interessanten Spezifikation: 4 der 5 Ärzte gruppieren sich zu 2, jeweils aus 2 Ärzten bestehenden Arzttypen. Die beiden Arzttypen unterscheiden sich im Ausmaß ihrer psychiatrischen Erfahrung. Die fachlich unerfahrenen Ärzte fühlen sich bei der psychologischen Beurteilung anfangs erheblich unsicher. Bei ihnen trägt die psychiatrisch besonders wichtige Information zur entscheidenden Veränderung des psychologischen Urteils, nicht zu einem besonders starken Anstieg der richtigen Diagnosen bei. Bei den psychiatrisch erfahreneren Ärzten ist der beschriebene Sachverhalt umgekehrt.

Die fachlich weniger erfahrenen Ärzte scheinen den Schwerpunkt ihrer Beurteilung auf die psychologische Beurteilung der Patienten gelegt zu haben. Die für eine psychiatrische Diagnose besonders wichtige Information wurde von ihnen weniger angemessen verarbeitet als von den fachlich erfahreneren Ärzten. Im Sinne der bisherigen Ausführungen kann vermutet werden, daß die fachlich weniger erfahrenen Ärzte eine nicht ausreichend klare Vorstellung vom psychiatrischen nosologischen System aufwiesen. Entsprechend waren sie nicht dazu in der Lage, die psychiatrisch besonders wichtige Information bewußt zu verarbeiten. Folglich tendierten sie - ähnlich wie Laien - zum "Psychologisieren", indem sie ihr Augenmerk hauptsächlich auf eine psychologische Beurteilung der Patienten richteten. Da für sie das psychologische Urteil besonders wichtig war, fühlten sie sich demgemäß auch anfänglich in ihrem psychologischen Urteil unsicher.

Aus den Ergebnissen zu den interindividuellen Urteilsdifferenzen lassen sich zusammenfassend zwei Schlußfolgerungen ziehen. Erstens ist aufgrund der z.T. großen interindividuellen Differenzen sowohl im psychiatrischen als auch psychologischen Urteil ein standardisiertes Trainingsprogramm für das bisher noch übliche Diagnostizieren zu fordern. Zweitens wird nach Abschaffung des klassischen nosologischen Systems die Anwendung eines modernen psychologischen Systems den psychiatrisch erfahrenen, im üblichen Diagnostizieren schon eingeschliffenen Psychiatern, beträchtliche Schwierigkei-

ten bereiten. Sie werden lernen müssen, psychologischer Information - ähnlich wie unerfahrene Psychiater - unsicher, damit aber auch flexibel gegenüberzustehen, um sie optimal aufzunehmen und verarbeiten zu können.

Auf die erwarteten beträchtlichen interindividuellen Unterschiede bei der Beurteilung explorativen Verhaltens von Psychiatern soll nicht näher eingegangen werden. Wichtiger erscheint die Tatsache, daß die Ärzte - entsprechend der Hypothese - das Verhalten bei der Exploration zur Erkrankung (Situation IV und V) klar vom Verhalten bei Kontaktaufnahme und biographischer Anamnese (Situation I und II) trennten. Interessanterweise wird - laut Urteil - die Bedeutung der Situationen für das explorative Verhalten im Zusammenhang mit bestimmten Patientengruppen gesehen. Dieser beurteilte Zusammenhang entspricht der realen Bedeutung der Situationen bei der Diagnose bestimmter Patientengruppen, die im ersten Teil der Untersuchung nachgewiesen wurde. So wird das Explorationsverhalten in den Situationen IV und V als insbesondere durch die Patientengruppen S und Z bestimmt eingeschätzt; diese Situationen liefern auch genau diejenige zusätzliche Information, die zur Diagnose von S und Z geeignet ist. Der unter diesen Bedingungen beobachtbare Explorationsstil scheint der typisch psychiatrische Explorationsstil zu sein. Offensichtlich handelt es sich um einen normierten Explorationsstil. Im Gegensatz dazu unterlag laut Urteil der Ärzte das explorative Verhalten in den beiden ersten Situationen sowie in Situation III gegenüber N einer geringeren Standardisierung. Die Ärzte konnten zwischen zwei Explorationsstilen wählen. Auch hier ist das Explorations-Verhalten von denjenigen Situationen und Patientengruppen bestimmt, die im realen psychiatrischen Explorationsvorgang eine besondere Einheit bilden; denn die Information aus Situation II und vermutlich insbesondere aus Situation III trägt zu einem besonders starken Anstieg der N-Diagnosen bei.

In dieser Untersuchung wurde der Versuch unternommen, den Einfluß von zunehmender und psychiatrisch relevanter werdender Information und von drei Patientengruppen auf das psychiatrische und psychologische Urteil von Psychiatern zu erfassen. Im Mittelpunkt des Interesses stand dabei der Verlaufsvergleich von psychologischem und psychiatrischem Urteil. Interindividuelle Urteilsdifferenzen zwischen den Psychiatern fanden ebenfalls Beachtung. Weiterhin wurde versucht, die Abhängigkeit explorativen Verhaltens von Explorationsthema und Patientengruppe zu überprüfen.

Die Psychiater, durch die die Daten zur Beantwortung der verschiedenen Fragen gewonnen wurden, übten in der Untersuchung eine doppelte Funktion aus. Sie fungierten nicht nur - im Sinne der Zielsetzung der Untersuchung - als Urteiler, sondern sie hatten auch die Exploration selbst durchzuführen. Diese zweite Tätigkeit beeinträchtigte möglicherweise die Urteilsfähigkeit. Dieses Problem hätte durch Fernseh- und Tonbandaufnahmen und spätere Darbietung der Aufnahmen zur Beurteilung umgangen werden können. Ein solches Vorgehen benötigt jedoch einen doppelten Zeitaufwand.

Kritikwürdig ist auch der Untersuchungsaufbau, wie schon im Abschnitt C: "Untersuchung" ausgeführt wurde. Diese Kritik betraf zum einen die dem Lateinischen Quadrat entsprechende Abfolge der Exploration durch die Psychiater. Diese geregelte Abfolge ist jedoch notwendig, um Abfolge-Effekte zu vereinheitlichen. Andere Untersuchungsansätze, wie Gesamtexploration eines jeden Patienten durch jeden Psychiater oder Gesamtexploration unterschiedlicher Patienten durch die Psychiater, bringen vermutlich noch größere Nachteile mit sich. Der zweite, schon genannte Kritikpunkt am Untersuchungsansatz zielt auf die Koppelung von Informationsumfang und -inhalt. Diese Koppelung war aus praktischen Erwägungen vorgenommen worden. In einer zur Zeit laufenden Untersuchung wird der voneinander unabhängige Einfluß beider Bedingungen auf das psychologische Urteil überprüft.

Es sei noch darauf hingewiesen, daß zusätzlich das verbale und mimische Verhalten der Patienten, insbesondere aber der explorierenden Psychiater hätte erfaßt und nach der Untersuchung analysiert werden sollen, um das explorative Verhalten der Psychiater umfassender in den Griff zu bekommen. Und weiterhin hätte man den Untersuchungsaufbau differenzierter gestalten können, mit dem Ziel, den Einfluß ganz spezifischer Informationsinhalte auf das Urteil der Psychiater zu erfassen. Dieser Ansatz hätte zur Aufklärung interindividueller Differenzen zwischen den Psychiatern beitragen können. Außerdem wären die aus einem solchen Ansatz resultierenden Ergebnisse von unmittelbarer, spezifischer praktischer Bedeutung für die Ausbildung in Explorationstechnik und üblicher psychiatrischer Diagnostik.

Die untersuchungskritischen Überlegungen werden durch folgende allgemeine Ausführungen abgeschlossen: Die Genauigkeit der Ergebnisse von empirisch-psychologischen Untersuchungen ist jeweils durch Versuchspersonen, deren testpsychologische Erfahrung, Meßinstrument und Wirkung des Untersuchungseingriffs in die soziale Wirklichkeit geprägt. Bei experimentellen Ansätzen, mit Hilfe deren (je vorläufige) psychologische

Gesetzmäßigkeiten überprüft werden, wird die zur Untersuchung notwendige soziale Wirklichkeit sogar manipulativ hergestellt, wobei die Realitätsnähe und somit die Bedeutung der ermittelten Gesetzmäßigkeiten vom Ausmaß der "Künstlichkeit" des experimentellen Ansatzes direkt abhängig ist.

Die relative "Natürlichkeit" der hier dargestellten Untersuchungssituation zu belegen, wurde mehrmals versucht. Die Untersuchungssituation entsprach - wie in der Einleitung aufgezeigt - in etwa der diagnostischen Untersuchungssituation in einer psychiatrischen Poliklinik, und folglich gelten die ermittelten Ergebnisse nur für diese Situation. Daß außerdem die Ergebnisse die untersuchten psychologischen Sachverhalte nicht unverzerrt widerspiegeln, ist aufgrund der genannten Überlegungen klar.

F. SCHLUSSBEMERKUNGEN

Unmittelbarer Gegenstand dieser psychologischen Untersuchung sind psychologisches und psychiatrisches Urteilen von Psychiatern über Patienten und ihr eigenes Verhalten in einer relativ "natürlichen" poliklinikähnlichen psychiatrisch-diagnostischen Untersuchungssituation. Die Arzt-Patient-Beziehung in dieser sozialen Situation stellt ein institutionalisiertes Interaktionssystem dar, in dem jeweils Arzt und Patient eindeutige Rollenmuster zugeschrieben sind.

Es soll nun kurz und grob-vereinfacht versucht werden anzudeuten - soweit das möglich ist -, inwiefern die hier ermittelten, eher unter denkpsychologischen Gesichtspunkten interpretierten Ergebnisse durch die Rollen von Patient und Arzt beeinflusst sind. Weiterhin werden neue Fragestellungen aufgezeigt.

Im Sinne der Pattern-Variablen von PARSONS (1951) ist die Rolle des Arztes durch Leistungsdenken, Universalismus, funktionale Spezifität, emotionale Neutralität und Kollektivitätsorientierung gekennzeichnet. Postuliert man diese Rollencharakteristika auch für die Rolle des Psychiaters, dann ist von ihm in der psychiatrisch-diagnostischen Poliklinik-Situation ein affektiv neutrales Verhalten zu erwarten, das darauf abzielt, möglichst schnell eine nicht emotional gefärbte psychiatrische Diagnose zu stellen, um den Patienten mit einer gezielten Therapie heilen zu können. Das der so definierten Rolle des Arztes entsprechende Verhalten scheint sich tatsächlich z.T. in den durchschnittlichen Ergebnissen bezüglich psychologischen und psychiatrischen Beurteilens

der Patienten niedergeschlagen zu haben. So scheinen Patienten bewertende Stellungnahmen kaum erfolgt zu sein, und vor allem scheint das psychiatrische Urteil im Verhältnis zum psychologischen bezüglich Bedeutsamkeit im Vordergrund gestanden zu haben, was insbesondere aufgrund der frühen Stabilisierung des psychologischen Urteils zu vermuten ist. Die Patienten scheinen in der psychiatrisch-diagnostischen Untersuchungssituation - entsprechend den Forderungen des klassischen psychiatrischen Diagnosesystems - lediglich als Träger psychiatrischer Symptome angesehen worden zu sein, an denen psychologische Eigenschaften nur im Zusammenhang mit der psychiatrischen Diagnose - und dort nur in geringem Maße - interessant waren. Diese Ausführungen implizieren, daß im Sinne der Rollencharakteristika Universalismus, Leistungsdenken und Spezifität das psychologische Urteil eine mindere Bedeutung hat, weil es zur geforderten spezifisch-psychiatrischen Diagnose wenig beiträgt (siehe die Erläuterung zur Funktion des psychologischen Urteils).

Diese für den Durchschnitt der Ärzte getroffenen Aussagen erfahren unter differentiellen Gesichtspunkten insofern eine Spezifizierung, als bei psychiatrisch unerfahrenen Ärzten ein eher gegenteiliges Ergebnis ermittelt wurde, d.h. diese Psychiater wichen in ihrem Urteilsverhalten von den Rollenerwartungen ab. Dieses Urteilsverhalten war als Tendenz zum "Psychologisieren" bezeichnet worden, das im Rahmen klassisch-psychiatrischen Diagnostizierens als unzureichend und für eine valide psychiatrische Diagnose als eher abträglich gilt. Eine Erläuterung dieses Urteilsverhaltens war schon versucht worden. Hier lassen sich nun neue psychologische Fragestellungen formulieren. So wäre untersuchenswert, ob außer dem Grad der psychiatrischen Erfahrung auch Einstellungs- und Persönlichkeitsvariablen der Psychiater die "psychologische" versus "psychiatrische" Diagnostikhaltung gegenüber den Patienten in der beschriebenen Untersuchungssituation beeinflussen und somit zu einer Änderung des Rollenverhaltens beitragen. Möglich ist z.B., daß die Einstellung, psychosoziale Faktoren seien bei der Genese auch endogen psychotischer Erkrankungen von allergrößter Bedeutung, einhergeht mit der Tendenz zum "Psychologisieren", was relativ späte Stabilisierung des psychologischen Urteils impliziert. Eine solche, immer auch gesellschaftskritische Haltung muß psychologische Faktoren am Patienten stärker berücksichtigen und kann sich nicht darauf beschränken, eine vorrangig als genetisch bedingt angesehene Erkrankung durch Überprüfung psychiatrischer Symptome zu erschließen. Dieser Verän-

derung des psychologisch-psychiatrischen Urteils könnte eine Abweichung des Rollenverhaltens in Bezug auf die Rollencharakteristika Universalismus, funktionale Spezifität und emotionale Neutralität und möglicherweise auch in Bezug auf Kollektivitätsorientiertheit und Leistungsdenken zugrunde liegen, was ein Abweichen von den genannten Rollenerwartungen insgesamt bedeutet. Wie weit mit diesem möglichen abweichenden Rollenverhalten vielleicht auch Veränderungen der Explorationsstile einhergehen, ist eine weitere überprüfenswerte Frage.

Die hier angerissenen Probleme fragen nach dem differentiellen Einfluß von Einstellungs- und Persönlichkeitsvariablen auf das Rollenverhalten des Arztes, das u.a. im psychologischen und psychiatrischen Urteil sowie im Explorationsverhalten seinen Ausdruck findet. Eine weitere interessante Frage ist die nach der situationsbedingten Änderung der Rolle von Arzt und Patient. Wichtigste Alternative zur diagnostischen Poliklinik-Situation ist die diagnostische Situation in der psychiatrischen Klinik, der sich auch stationär aufgenommene Patienten zu unterziehen haben. Auch in dieser sozialen Situation laufen diagnostische Prozesse, insbesondere zu Anfang einer Aufnahme ab, die aber sehr bald durch therapeutische Komponenten erweitert werden. Die Rolle des Arztes dürfte sich - zumindest nach PARSONS (1951) - unter diesen Bedingungen nur wenig ändern. Einen wichtigen Wandel erfährt jedoch wahrscheinlich die Rolle des Patienten, dem sich - ohne jetzt näher auf die Rollencharakteristika des Patienten laut PARSONS (1951) einzugehen - im Verhältnis zu der auf die soziale Beziehung zum Arzt eingeschränkten Poliklinik-Situation nun eine Vielzahl möglicher sozialer Interaktionen bietet (Mitpatienten, Klinikpersonal, Besuch usw.). Es ist anzunehmen, daß der Patient unter diesen sozialen Bedingungen Verhalten äußert, das im Verhältnis zu dem in der klar standardisierten, auf eine psychiatrische Diagnose zentrierten und auf den Arzt beschränkten sozialen Poliklinik-Situation von größerer psychologischer Bedeutung ist. Damit sind möglicherweise für den auf der Station beobachtenden Arzt gute Voraussetzungen gegeben, sich ein nur langsam stabilisierendes psychologisches Bild vom Patienten zu bilden. Die Anzahl psychiatrischer Diagnosewechsel dürfte unter diesen Bedingungen dagegen stark abnehmen.

Der Einfluß der Klinik-Situation auf das Verhalten des Psychiaters gegenüber den Patienten bedarf - ebenfalls wie die obige Fragestellung - einer Untersuchung.

Die beschriebenen sozialen Situationen lassen sich natürlich weitaus differenzierter analysieren und entsprechend weitere Fragestellungen ableiten. Das jedoch ist nicht - wie anfangs erwähnt - beabsichtigt. Abschließend ist es wichtiger, nun noch die Be-

deutung der hier durchgeführten Untersuchung und der angerissenen weiteren Fragestellungen kurz abzuklären.

Durch die Untersuchung wurden einige Gesetzmäßigkeiten vorläufig anerkannt; im Zusammenhang mit der Beantwortung der weiteren Fragestellungen kann eine Theorie über psychologisches und psychiatrisches Urteilen und Explorationsverhalten von Psychiatern erstellt werden. Abgesehen davon, daß damit psychologische Gegebenheiten erklärt werden, implizieren die Gesetze einer solchen Theorie technische Anwendungsmöglichkeiten, d.h. beabsichtigte Herstellung von Bedingungen, um bestimmte Wirkungen zu erreichen. Die Kenntnis dieser Möglichkeiten kann in diesem Fall zum Teil nicht und darf auch aus ethischen Gründen hier prinzipiell nicht angewendet werden; vielmehr muß sie zur Aufklärung der Psychiater über ihr eigenes Urteilen und exploratives Verhalten herangezogen werden. Insbesondere muß auf solche empirischen Gesetzmäßigkeiten aufmerksam gemacht werden, die im Sinne üblicher psychiatrischer Diagnosefindung nicht wünschenswert sind (Abhängigkeit der Diagnosen von jeweiligen Explorationsinhalten und weitere erwähnte Beispiele). Somit wird also die Überwindung einiger empirischer Gesetzmäßigkeiten angezielt, die Untersuchung verfolgt insofern auch kritische Interessen. Die Entscheidung für Änderung empirischer Gesetzmäßigkeiten basiert hier natürlich auf der Überlegung, daß die entsprechenden diagnostischen Effekte - in welchem diagnostischen Sinne (z.B. Diagnosesystem, Ziel der Diagnostik) auch immer - nicht wünschenswert sind. Hier zeichnet sich ab, daß die weitere Diskussion über die Bedeutung der ermittelten empirischen Gesetzmäßigkeiten einmündet in die Fragen, die durch die Gesetzmäßigkeiten selbst natürlich nicht mehr gelöst werden können. Hierher gehört insbesondere die Frage nach der gesellschaftlichen Bedeutsamkeit psychiatrischen und psychologischen Diagnostizierens und Urteilens und damit die Frage nach der gesellschaftlichen Bedeutung und Funktion der Psychiatrie selbst; entsprechende Fragen waren bei der Diskussion der Psychiater-Rolle ausgeklammert worden.

Dazu gibt es verschiedene Ansätze in der medizin-soziologischen Literatur, die - den hier aufgezeigten Rahmen meist sprengend - psychiatrische Kliniken insgesamt als soziale Systeme einer Analyse unterziehen.

Literaturverzeichnis

- ANDERSON, N.H. Integration theory and attitude change.
Psychol. Rev., 1971, 78 (3), 171-206
- ARTHUR, R.J.;
GUNDERSON, E.K. The prediction of diagnosis and disposition in naval
hospitals.
J. clin. Psychol., 1966, 22, 259-264
- ASCH, S.E. Forming impressions of personality.
J. abnorm. soc. Psychol., 1964, 41, 258-290
- BEACH, L.;
WERTHEIMER, M. A free response approach to the study of person cognition.
J. abnorm. soc. Psychol., 1961, 62, 367-374
- BECHTOLDT, H.P. An empirical study of the factor analysis stability hypothesis.
Psychometrika, 1961, 26, 405-432
- BELLMAN, R.R. Introduction to matrix analysis.
N.Y., 1960
- BIERI, J.;
ATKINS, A.L.;
BRIAR, S.;
LEAMAN, R.L.;
MILLER, H.;
TRIPODI, T. Clinical and social judgment: The discrimination of be-
havioral information.
N.Y., 1966
- BRUNER, J.S.;
TAGIURI, R. The perception of people.
In: Lindzey, G. (ed.), Handbook of social psychology
Vol. II, Cambridge, 1954, 601-633
- BRUNER, J.S.;
GOODNOW, J.J.;
AUSTIN, G.A. A study of thinking.
N.Y., 1956
- BRUNER, J.S.;
SHAPIRO, D.;
TAGIURI, R. The meaning of traits in isolation and in combination.
In: Tagiuri, R. and Petrullo, L. (eds.), Person per-
ception and interpersonal behavior. Stanford, 1958,
277-288
- BRUNSWIK, E. Representative design and probabilistic theory in a
functional psychology.
Psychol. Rev., 1955, 62, 193-217
- BRUNSWIK, E. Perception and the representative design of psychological
experiments.
Berkeley, 1956
- CAMPBELL, D.T.;
FISKE, D.W. Convergent and discriminant validation by the multitrait-
multimethod matrix.
Psychol. Bull., 1959, 56, 81-100

- CATTELL, R.B. Personality and motivation structure and measurement.
N.Y., 1957
- CATTELL, R.B.;
SCHEIER, J.H. The meaning and measurement of neurotizism and anxiety.
N.Y., 1961
- COHEN, R. Eine Untersuchung zur diagnostischen Verarbeitung wider-
sprüchlicher Informationen.
Psychol. Forsch., 1967, 30, 211-225
- COHEN, R. Systematische Tendenzen bei Persönlichkeitsbeurteilungen.
Bern, 1969
- CROCKETT, W.A. Cognitive complexity and impression formation.
In: Maher, B.A. (ed.), Progress in experimental personality
research. Vol. II, N.Y., 1965
- CRONBACH, L.J. Processes affecting scores on the "understanding of others"
and "assumed similarity".
Psychol. Bull., 1955, 52, 177-193
- CRONBACH, L.J. Proposals leading to analytic treatment of social perception
scores.
In: Tagiuri, R. and Petrullo, L. (eds.), Person perception
and interpersonal behavior. Stanford, 1958, 353-379
- DEROGATIS, L.R.;
LIPMAN, R.S.;
COVI, L.;
RICKELS, K.;
UHLENHUTH, E.H. Dimensions of outpatient neurotic pathology: Comparison
of a clinical versus an empirical assessment.
J. consult. clin. Psychol., 1970, 34 (2), 164-171
- DORNBUSCH, S.M.;
HASTORF, A.H.;
RICHARDSON, S.A. The perceiver and the perceived. Their relative influence
on the categories of interpersonal cognition.
J. Pers. soc. Psychol., 1965, 1, 434-440
- ECKART, C.;
YOUNG, G. The approximation of one matrix by another of lower
rank.
Psychometrika, 1936, 1, 211-218
- FELIPE, A.J. Evaluative versus descriptive consistency in trait inference.
J. Pers. soc. Psychol., 1970, 16 (4), 627-638
- FISCHER, G.;
ROPPERT, J. Bemerkungen zu einem Verfahren der Transformationsana-
lyse.
Arch. ges. Psychol., 1964, 116, 98-100
- FRANK, R. Klinische Urteilsbildung.
Unveröff. Dipl.-Arbeit am Psychol. Inst. Giessen, 1970

- GEBHARDT, F. Über Ähnlichkeit von Faktorenmatrizen.
Psychol. Beiträge, 1967, 10, 591-599
- GITIN, S.R. A dimensional analysis of manual expression.
J. Pers. soc. Psychol., 1970, 15 (3), 271-277
- GOLDBERG, L.R. The reliability of clinicians' judgments: A multitrait-
multimethod approach.
J. consult. Psychol., 1966, 30, 199-206
- GRAUMANN, C.F. Eigenschaften als Problem der Persönlichkeitsforschung.
In: Lersch, P. et al. (Hrsg.): Persönlichkeitsforschung.
Hndb. Psychologie, Bd. 4, Göttingen, 1960, 87-154
- HATFIELD, J.S.;
SHAPIRO, A.M. Dimensions among interview ratings of men: An analysis
of their meaning and personality inventory correlates.
J.clin.Psychol., 1965, 21 (4), 150-160
- HAYS, W.L. An approach to the study of trait implication and trait
similarity.
In: Tagiuri, R. and Petrullo, L. (eds.), Person perception
and interpersonal behavior. Stanford, 1958, 289-299
- HOFSTÄTTER, P.R. Einführung in die Sozialpsychologie.
Stuttgart, 1963
- KAMINSKI, G. Das Bild vom Anderem.
Berlin, 1959
- KEIL, W. Implizite Persönlichkeitstheorie und Urteile über Andere.
Zeitschr. Sozialpsychol., 1970, 1, 351-360
- KELLY, G.A. The psychology of personal constructs.
N.Y., Norton, 1955 (2 Bde)
- KOLTUR, B.B. Some characteristics of intrajudge trait intercorrelations.
Psychol. Monogr., 1962, 76, Whole No. 552
- LEVIN, J. Three-mode factor analysis.
Psychol. Bull., 1965, 64, 442-452
- LOEHLIN, J.C. Word meanings and self-descriptions.
J. Pers. soc. Psychol., 1967, 5, 107-110
- LORR, M.;
KLETT, C.J.;
McNAIR, D.M. Syndromes of psychosis.
Oxford, 1963

- MEEHL, P.E. The cognitive activity of the clinician.
Amer. Psychologist, 1960, 15, 19-27
- MILLER, H.;
TRIPODI, T. Clinical and social judgment: The discrimination of
behavioral information.
N.Y., 1966
- Münchener Symposium über Faktorenanalyse.
Berlin, 1968
- NEWCOMB, T.M. The prediction of interpersonal attraction.
Amer. Psychologist, 1956, 11, 575-586
- NORMAN, W.T.;
GOLDBERG, L.R. Raters, ratees, and randomness in personality structure.
J. Pers. soc. Psychol., 1966, 4, 681-691
- OSGOOD, C.E.;
SUCI, G.J.;
TANNENBAUM, P.H. The measurement of meaning.
Illinois, 1957
- PARSONS, T. The social system.
Glencoe, 1951
- PASSINI, F.T.;
NORMAN, W.T. A universal conception of personality structure?
J. Pers. soc. Psychol., 1966, 4, 44-49
- PEABODY, D. Trait inferences: Evaluative and descriptive aspects.
J. Pers. soc. Psychol., Monograph 1967, 7, 4
- PEABODY, D. Evaluative and descriptive aspects in personality per-
ception: A reappraisal.
J. Pers. soc. Psychol., 1970, 16 (4), 639-646
- PETERSON, D.R. Scope and generality of verbally defined personality
factors.
Psychol. Rev., 1965, 72, 48-59
- RANABAUER, W. Psychiatrische Diagnosen unter psychologischen Ge-
sichtspunkten.
Nervenarzt, 1968, 39 (5), 205-213
- REDLICH, F.C.;
FREEDMAN, D.X. The theory and practice of psychiatry.
N.Y., 1966
- ROSENBERG, S.;
OLSHAN, K. Evaluative and descriptive aspects in personality per-
ception.
J. Pers. soc. Psychol., 1970, 16 (4), 619-626
- SARBIN, T.R.;
TAFT, R.;
BAILEY, D.E. Clinical inference and cognitive theory.
N.Y., 1960

- SCHEFF, T.J. Decision rules, types of error, and their consequences in medical diagnosis.
Behav. Scie., 1963, 8, 97-107
- SHRANGER, S. Cognitive differentiation and the impression-formation process.
J. Pers., 1967, 35, 402-414
- TAGIURI, R. Person perception.
In: Hndb. Social Psychology, Vol. III, 395-449, 1969²
- THORNDIKE, E.L. Fundamental theorems in judging men.
J. appl. Psychol., 1918, 2, 67-76
- TODD, F.J.;
RAPPOPART, L. A cognitive structure approach to person perception.
J. abnorm.soc. Psychol., 1964, 68, 469-478
- TUCKER, L.R. The extension of factor analysis to three-dimensional matrices.
In: Frederiksen, N. and Gulliksen, H. (eds.), Contributions to mathematical psychology. N.Y., 1964
- TUCKER, L.R. Some mathematical notes on three-mode factor analysis.
Psychometrika, 1966, 31 (3), 279-311
- TUCKER, L.R.;
MESSICK, S. An individual differences model for multidimensional scaling.
Psychometrika, 1963, 28, 333-367
- ÜBERLA, K. Faktorenanalyse.
Berlin, 1968
- WARD, C.H.;
BECK, A.T.;
MENDELSON, M.;
MOCK, J.E.;
ERBAUGH, J.K. Arch. gen. Psychiat., 1962, 7, 198-205
- WARR, P.B.;
KNAPPER, C. The perception of people and events.
London, 1968
- ZUBIN, J. Classification of the behavior disorders.
Ann. Rev. Psychol., 1967, 18, 373-406

Anhang I

PATIENTEN - POLARITÄTEN-PROFIL

Beurteiler:

Objekt der Beurteilung:

Nach Explorationsbereich: I III V

freundlich-ausgeglichen	1--2--3--4--5--6--7	mürrisch-gereizt
weitschweifig	1--2--3--4--5--6--7	zielgerichtet
verschlossen	1--2--3--4--5--6--7	offen
differenziert	1--2--3--4--5--6--7	undifferenziert
antriebsgehemmt	1--2--3--4--5--6--7	antriebsgesteigert
beharrlich	1--2--3--4--5--6--7	ablenkbar
gelöst	1--2--3--4--5--6--7	gespannt
traurig	1--2--3--4--5--6--7	heiter
mißtrauisch	1--2--3--4--5--6--7	arglos
sich unterordnend	1--2--3--4--5--6--7	führend
stumpf	1--2--3--4--5--6--7	sensibel
sprunghaft	1--2--3--4--5--6--7	geordnet
warmherzig	1--2--3--4--5--6--7	gefühlslarm
selbstbewußt	1--2--3--4--5--6--7	selbstunsicher
aggressiv	1--2--3--4--5--6--7	ängstlich-zurückhaltend
interessant	1--2--3--4--5--6--7	langweilig
resonanzunfähig	1--2--3--4--5--6--7	emotional gut ansprechbar
ratlos	1--2--3--4--5--6--7	entschieden
unglaublich	1--2--3--4--5--6--7	ehrlich
störbar	1--2--3--4--5--6--7	unbeeindruckbar
sympathisch	1--2--3--4--5--6--7	unsympathisch

Entscheidung : sehr leicht 1--2--3--4--5 sehr schwer

Vermutete Diagnose :

Sicherheit der Diagnose : sicher 1--2--3--4--5 sehr unsicher

ÄRZTE - POLARITÄTEN-PROFIL

Beurteiler:

Objekt der Beurteilung:

aktiv	1--2--3--4--5--6--7	passiv
rational	1--2--3--4--5--6--7	gefühlsmäßig
optimistisch	1--2--3--4--5--6--7	pessimistisch
geduldig	1--2--3--4--5--6--7	ungeduldig
im Verhalten ausdrucksvoll	1--2--3--4--5--6--7	im Verhalten farblos
weitschweifig	1--2--3--4--5--6--7	zielgerichtet
selbstgefällig	1--2--3--4--5--6--7	selbstkritisch
emotional mitschwingend	1--2--3--4--5--6--7	emotional starr
irritierbar	1--2--3--4--5--6--7	gelassen
still	1--2--3--4--5--6--7	lebhaft
arglos	1--2--3--4--5--6--7	mißtrauisch
streng	1--2--3--4--5--6--7	gewährend
originell vorgehend	1--2--3--4--5--6--7	schematisch vorgehend
selbstbezogen	1--2--3--4--5--6--7	sachlich
langsam	1--2--3--4--5--6--7	schnell
kontrolliert	1--2--3--4--5--6--7	unbefangen
freundlich-aufmunternd	1--2--3--4--5--6--7	unbeteiligt-abwartend
beharrend	1--2--3--4--5--6--7	schnell aufgebend
interessiert	1--2--3--4--5--6--7	gelangweilt
gespannt	1--2--3--4--5--6--7	gelöst
wechselnd	1--2--3--4--5--6--7	gleichbleibend

Anhang III

Arzt- faktor	Patienten- faktor	Skalen- faktor	1	2	3	4
1	1		-12,03	6,09	-2,57	-6,77
	2		-5,56	6,64	1,57	-3,92
	3		-,92	-,37	,32	7,90
	4		-,73	-4,72	-3,23	-3,06
2	1		8,45	3,23	4,21	,08
	2		1,12	-8,10	-1,42	4,83
	3		-,10	4,58	-,64	-,70
	4		-,09	5,46	2,42	4,94

Kernmatrix aus Urteil 1

Arzt- faktor	Patienten- faktor	Skalen- faktor	1	2	3	4
1	1		10,17	4,65	-3,90	-5,17
	2		3,78	5,87	1,40	-5,62
	3		-4,23	6,19	2,92	10,10
	4		1,79	3,11	-6,69	6,71
2	1		2,92	1,79	4,30	10,16
	2		-8,28	-7,52	-4,46	2,03
	3		5,59	-4,31	-4,06	6,06
	4		2,69	-2,56	1,78	-6,19

Kernmatrix aus Urteil 3

Lebenslauf

22.9.1942	Geburt in Bielitz/Oberschlesien als Sohn von Ernst Leichner und dessen Ehefrau Susanne
Ostern 1949	Einschulung in die Volksschule in Wilhelmsdorf/Taunus
Ostern 1955	Besuch des Aufbaugymnasiums in Usingen/Taunus
Ostern 1962	Abitur an derselben Schule
SS 1962	Beginn des Studiums der Betriebswirtschaft an der Universität Frankfurt
WS 1962/63	Wechsel des Studienfaches und Anfang des Psychologiestudiums an derselben Universität
SS 1964	Fortführung des Psychologiestudiums an der Universität Giessen
Frühjahr 1965	Vordiplomprüfung in Psychologie an der Universität Giessen
Frühjahr 1968	Hauptdiplomprüfung in Psychologie an derselben Universität
Mai 1968	Beginn einer klinisch-diagnostischen Tätigkeit an der Neuropsychiatrischen Klinik der Universität Giessen
April 1970	Aufnahme der Tätigkeit als wissenschaftlicher Assistent (m.d.V.b.) (jetzt wissenschaftlicher Angestellter) am Psychologischen Institut (jetzt Fachbereich 06 Psychologie) an der Universität Giessen
	Bisher geleitete Veranstaltungen: Theoretische Aspekte der Diagnostik, Diagnostisches Praktikum, Sozial- und denkpsychologische Aspekte der Personwahrnehmung, Denken und Intelligenz, Denken und Sprache, Zentrale Denk- und Sprachstörungen.
	Die Forschungsinteressen und -ansätze stimmen mit den Themen der Veranstaltungen überein.
	Zukünftige Lehrinteressen: Methodenlehre, psychologische Meßmethoden

